

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:08:25 ; Search time 411.58 Seconds  
(without alignments)  
4.500 Million cell updates/sec

Title: US-09-724-406-24

Perfect score: 90

Sequence: 1 DPPYGNPHYAMDY 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_16.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_unclassified.\*

13: sp\_vertebrate.\*

14: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	53.3	470	10 Q9M301	Q9M301 arabidopsis
2	46	51.1	343	2 Q9KJN9	Q9KJN9 myxococcus
3	46	51.1	919	4 Q9NQS7	Q9NQS7 homo sapien
4	45	50.0	761	3 Q9H599	Q9H599 ashbya goss
5	44	48.9	527	13 Q9DGB4	Q9DGB4 xenopus lae
6	43.5	48.3	2573	5 Q9G185	Q9G185 plasmodium
7	43	47.8	276	4 Q9NWA9	Q9NWA9 homo sapien
8	43	47.8	329	1 O58076	O58076 pyrococcus
9	43	47.8	381	10 Q41556	Q41556 triticum ae
10	43	47.8	484	4 Q9H7L0	Q9H7L0 homo sapien
11	43	47.8	614	4 Q9NW41	Q9NW41 homo sapien
12	43	47.8	855	4 Q9NYK9	Q9NYK9 homo sapien
13	43	47.8	878	4 Q9UFU7	Q9UFU7 homo sapien
14	43	47.8	938	11 Q9JL72	Q9JL72 mus musculu
15	43	47.8	1080	2 O84688	O84688 chlamydia t
16	43	47.8	2809	5 O61230	O61230 lytechinus
17	42	46.7	497	2 Q9KBN0	Q9KBN0 bacillus ha
18	42	46.7	553	5 Q9GR06	Q9GR06 leishmania
19	42	46.7	860	2 Q9L428	Q9L428 alcaligenes

20	42	46.7	903	6 Q9GLY5	Q9GLY5 oryctolagus
21	42	46.7	973	2 O86728	O86728 streptomyce
22	41.5	46.1	363	2 Q9PIS1	Q9PIS1 campylobact
23	41	45.6	26	10 Q9S8A1	Q9S8A1 pinus monti
24	41	45.6	127	5 O17467	O17467 schistosoma
25	41	45.6	155	2 O33696	O33696 streptococc
26	41	45.6	218	2 Q9Z505	Q9Z505 streptococc
27	41	45.6	222	2 Q9S2D6	Q9S2D6 streptomyce
28	41	45.6	261	2 Q9K6A4	Q9K6A4 bacillus ha
29	41	45.6	357	7 Q30220	Q30220 ateles belz
30	41	45.6	359	4 Q06416	Q06416 homo sapien
31	41	45.6	612	8 Q9T9D0	Q9T9D0 paralichthy
32	41	45.6	950	5 Q20733	Q20733 caenorhabdi
33	41	45.6	973	10 Q9S9Q8	Q9S9Q8 arabidopsis
34	41	45.6	1194	5 Q9W485	Q9W485 drosophila
35	41	45.6	1413	5 Q9VK08	Q9VK08 drosophila
36	41	45.6	1590	2 Q55263	Q55263 streptococc
37	41	45.6	1590	2 Q59983	Q59983 streptococc
38	41	45.6	2197	5 Q96296	Q96296 plasmodium
39	41	45.6	2379	4 Q9Y493	Q9Y493 homo sapien
40	41	45.6	2710	5 Q9XZB8	Q9XZB8 plasmodium
41	40.5	45.0	407	10 O82036	O82036 petroselinu
42	40.5	45.0	420	10 O82037	O82037 petroselinu
43	40	44.4	171	1 Q59551	Q59551 pyrococcus
44	40	44.4	204	10 Q38895	Q38895 arabidopsis
45	40	44.4	246	10 Q9SV94	Q9SV94 arabidopsis

#### ALIGNMENTS

RESULT 1  
Q9M301 ID Q9M301 PRELIMINARY; PRT; 470 AA.  
AC Q9M301;

DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE HYPOTHETICAL 53.2 KDA PROTEIN.  
GN T21J18\_90.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsi

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,

RA Meves H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,

RA Salanoubat M.;

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL132963; CAB87910.1; -

DR InterPro; IPR001675; -

DR Pfam; PF00777; Glyco-transf\_29; 1.

DR Hypothetical protein.

SQ SEQUENCE 470 AA; 53188 MW; 0D0B8EA4CB52C656 CRC64;

Query Match 53.3%; Score 48; DB 10; Length 470;

Best Local Similarity 63.6%; Pred. No. 13;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 YGNPHYAMDY 14

Db 83 YGLPYIVLDY 93

RESULT 2

Q9KJN9

ID Q9KJN9 PRELIMINARY; PRT; 343 AA.  
 AC Q9KJN9;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE HYPOTHETICAL 36.6 KDA PROTEIN.  
 OS Myxococcus xanthus.  
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
 OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.  
 OX NCBI\_TaxID=34;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DZ2;  
 RX MEDLINE=20032045; PubMed=10564511;  
 RA Cho K., Zusman D.R.;  
 RT "Sporulation timing in Myxococcus xanthus is controlled by the espAB  
 locus.";  
 RL Mol. Microbiol. 34:714-725(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DZ2;  
 RA Cho K., Zusman D.R.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF163841; AAF87928.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 343 AA; 36580 MW; 16B89C41B1023979 CRC64;

Query Match 51.1%; Score 46; DB 2; Length 343;  
 Best Local Similarity 77.8%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PPGNPHY 10  
 |||||  
 Db 281 PPYGEPPY 289

RESULT 3

ID Q9NQS7 PRELIMINARY; PRT; 919 AA.  
 AC Q9NQS7;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE INNER CENTROMERE PROTEIN INCENP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Adams R.R., Eckley D.M., Mackay A., Earnshaw W.C.;  
 RT "Characterization and expression of the human INCENP gene.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF282265; AAF87584.1; -  
 SQ SEQUENCE 919 AA; 105536 MW; 8F93024E411087C6 CRC64;

Query Match 51.1%; Score 46; DB 4; Length 919;  
 Best Local Similarity 66.7%; Pred. No. 55;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 PPGNPHYAMD 13  
 |||||  
 Db 815 PPQINPNHYGMD 826

RESULT 4

ID Q9HF59 PRELIMINARY; PRT; 761 AA.  
 AC Q9HF59;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE CDC24.  
 GN Ashbya gossypii (Eremothecium gossypii).  
 OS Ashbya gossypii (Eremothecium gossypii).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Eremotheciaceae; Eremothecium.  
 OX NCBI\_TaxID=33169;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wendland J., Philippsen P.;  
 RT "Hyphal morphogenesis is controlled by multiple Rho-protein modules in  
 the filamentous ascomycete Ashbya gossypii.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF210626; AAG41244.1; -  
 SQ SEQUENCE 761 AA; 86025 MW; C7334F46A06BB334 CRC64;

Query Match 50.0%; Score 45; DB 3; Length 761;  
 Best Local Similarity 77.8%; Pred. No. 64;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PYGNPHYA 11  
 |||||  
 Db 628 PYGHPHYSA 636

RESULT 5

ID Q9DGB4 PRELIMINARY; PRT; 527 AA.  
 AC Q9DGB4;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE DNA POLYMERASE EPSILON SUBUNIT B.  
 GN POLE2.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OVARY;  
 RA Waga S., Sugino A.;  
 RT "Xenopus DNA polymerase epsilon subunit B.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB048257; BAB12726.1; -  
 SQ SEQUENCE 527 AA; 59688 MW; CF25EE4FB9C711C7 CRC64;

Query Match 48.9%; Score 44; DB 13; Length 527;  
 Best Local Similarity 53.8%; Pred. No. 61;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PPGNPHYAMDY 14  
 ||:|:|:|  
 Db 455 PPFSVPVWAYDY 467

RESULT 6

ID O96185 PRELIMINARY; PRT; 2573 AA.  
 AC O96185;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE HYPOTHETICAL 308.1 KDA PROTEIN.  
 GN PFB0460C.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Perle M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL; AE001396; AAC11881.1; -
KW Hypothetical protein.
SQ SEQUENCE 2573 AA; 308142 MW; 49E0D7E8967066CF CRC64;

Query Match 48.3%; Score 43.5; DB 5; Length 2573;
Best Local Similarity 72.7%; Pred. No. 4.2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 DP-PYGNPHYY 10
II I I I I I I
DB 1436 DPEAYNPHYY 1446

RESULT 7
Q9NWA9 PRELIMINARY; PRT; 276 AA.
AC Q9NWA9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 29.6 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Makamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001032; BAA91474.1; -
DR InterPro; IPR000286; -
DR Pfam; PF00850; Hist_deacetyl1; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 276 AA; 29622 MW; 8C96C7C5B41CD3A1 CRC64;

Query Match 47.8%; Score 43; DB 4; Length 276;
Best Local Similarity 63.6%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDPYGNPHYY 11
III I I I I I
DB 90 DPMGDPYLA 100

RESULT 8
O58076 PRELIMINARY; PRT; 329 AA.
ID O58076
AC O58076
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 37.2 KDA PROTEIN PH0338.
GN PH0338.
OS Pyrococcus horikoshii.

Pyrococcus horikoshii.

```

```

OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000002; BAA29412.1; -
DR InterPro; IPR000051; -
DR InterPro; IPR000241; -
DR InterPro; IPR002052; -
DR Pfam; PF01170; UPF0020; 1.
DR PROSITE; PS00092; N6_MTAASE; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 329 AA; 37237 MW; 21485A787DB28294 CRC64;

Query Match 47.8%; Score 43; DB 1; Length 329;
Best Local Similarity 61.5%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDPYGNPHYYAMD 13
II I I I I I I
DB 255 DDPYGNSTTLPMD 267

RESULT 9
Q41556 PRELIMINARY; PRT; 381 AA.
ID Q41556
AC Q41556
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TRANSCRIPTION FACTOR HBP-1A.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae;
OC Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HORSHIRIKOMUGI;
RA Nakayama T.;
RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94193693; PubMed=8144592;
RA Mikami K., Sakamoto A., Iwabuchi M.;
RT "The HBP-1 family of wheat basic/leucine zipper proteins interacts
RT with overlapping cis-acting hexamer motifs of plant histone genes.";
RL J. Biol. Chem. 269:9974-9985(1994).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: TO THE BZIP FAMILY.
DR EMBL; D12920; BAA02304.1; -
DR Mendel; 15812; Triae; 2979; 15812.
DR InterPro; IPR001871; -
DR Pfam; PF00170; bZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR SMART; SM00338; BRLZ; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 381 AA; 40683 MW; 448635628C15651F CRC64;

Query Match 47.8%; Score 43; DB 10; Length 381;

```

Best Local Similarity 61.5%; Pred. No. 61;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 PPYGNPHYA 14  
||||| |||  
Db 77 PPYGTPTPTMY 89

## RESULT 10

Q9H7L0 Q9H7L0 PRELIMINARY; PRT; 484 AA.  
AC Q9H7L0;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE FLJ00062 PROTEIN (FRAGMENT).  
GN FLJ00062.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RA Ohara O., Nagase T., Kikuno R., Okumura K.;  
RT "The nucleotide sequence of a long cDNA clone isolated from human  
RT spleen";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024469; BAB15759.1; -;  
FT NON\_TER 1  
SQ SEQUENCE 484 AA; 52327 MW; 1B1B0B363436A800 CRC64;

Query Match 47.8%; Score 43; DB 4; Length 484;  
Best Local Similarity 63.8%; Pred. No. 80;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDPYGNPHYA 11  
||||| |||  
Db 298 DPPMGDPYLA 308

## RESULT 11

Q9NW41 Q9NW41 PRELIMINARY; PRT; 614 AA.  
AC Q9NW41;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE CNA FLJ10328 FIS, CLONE NT2RM2000588, WEAKLY SIMILAR TO HISTONE  
DE DEACETYLASE HDAL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.;  
RP "NEDO human cDNA sequencing project";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK001190; BAA91545.1; -;  
DR InterPro; IPR000286; -;  
DR Pfam; PF00850; Hist.deacetyl1; 1.  
DR PRINTS; PR01270; HDASUPER.  
SQ SEQUENCE 614 AA; 66173 MW; 36C82D314014F211 CRC64;

Query Match 47.8%; Score 43; DB 4; Length 614;  
Best Local Similarity 63.6%; Pred. No. 1e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDPYGNPHYA 11  
||||| |||  
Db 428 DPPMGDPYLA 438

## RESULT 12

Q9NYK9 Q9NYK9 PRELIMINARY; PRT; 855 AA.  
AC Q9NYK9;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HISTONE DEACETYLASE 7.  
GN HDAC7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=CERVICAL CARCINOMA;  
RA Li S., Fischle W., Verdin E., Walsh M.J.;  
RT "A novel class II HDAC is associated with the transcriptional  
RT homeodomain repressor CCAAT displacement protein";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF239243; AAF63491.1; -;  
DR InterPro; IPR000286; -;  
DR Pfam; PF00850; Hist.deacetyl1; 1.  
DR PRINTS; PR01270; HDASUPER.  
SQ SEQUENCE 855 AA; 92363 MW; 3157426AE1913DB6 CRC64;

Query Match 47.8%; Score 43; DB 4; Length 855;  
Best Local Similarity 63.6%; Pred. No. 1.5e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDPYGNPHYA 11  
||||| |||  
Db 729 DPPMGDPYLA 739

## RESULT 13

Q9UFU7 Q9UFU7 PRELIMINARY; PRT; 878 AA.  
AC Q9UFU7;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 94.3 KDA PROTEIN (FRAGMENT).  
GN DKFZP586J0917.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=UTERUS;  
RA Koehrer K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL117455; CAB55935.1; -;  
DR InterPro; IPR000286; -;  
DR InterPro; IPR001781; -;  
DR Pfam; PF00850; Hist.deacetyl1; 1.  
DR PRINTS; PR01270; HDASUPER.  
DR Probom; PD000094; -; 2.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 878 AA; 94329 MW; 26882E693960E10D CRC64;



Query Match 47.8%; Score 43; DB 4; Length 878;  
 Best Local Similarity 63.6%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DPPYGNPHYA 11  
 ||| | | | |  
 DB 692 DPPMGDPYLA 702

## RESULT 14

Q9JL72 PRELIMINARY; PRT; 938 AA.  
 AC Q9JL72;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HISTONE DEACETYLASE 7.  
 GN HDAC7  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B6:C57BL/6;  
 RX MEDLINE=20107033; PubMed=10640276;  
 RT "Isolation of a novel histone deacetylase reveals that class I and  
 class II deacetylases promote SMRT-mediated repression.";  
 RL Genes Dev. 14:55-66(2000).  
 DR ENBL; AF207749; AAF31419.1; -.  
 DR InterPro; IPR000286; -.  
 DR Pfam; PF00850; Hist\_deacetyl; 1.  
 DR PRINTS; PRO1270; HDASUPER.  
 SQ SEQUENCE 938 AA; 101318 MW; A282A36EBC840897 CRC64;

Query Match 47.8%; Score 43; DB 11; Length 938;  
 Best Local Similarity 63.6%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DPPYGNPHYA 11  
 ||| | | | |  
 DB 753 DPPMGDPYLA 763

## RESULT 15

O84688 PRELIMINARY; PRT; 1080 AA.  
 AC O84688;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE PB2-TRANGLYCOLASE/TRANSEPTIDASE.  
 GN PB2.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/CX;  
 RX MEDLINE=9900809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 DR EMBL; AE001338; AAC68277.1; -.  
 DR InterPro; IPR001460; -.  
 DR Pfam; PF00905; Transpeptidase; 1.  
 SQ SEQUENCE 1080 AA; 123959 MW; EE647FA87F1FFEDD CRC64;

Query Match 47.8%; Score 43; DB 2; Length 1080;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DPPYGNPHY 9  
 :|| |||:  
 DB 662 EPPLGNPHY 670

Search completed: June 28, 2001, 16:08:26  
 Job time: 954 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:14:37 ; Search time 362.28 seconds  
(without alignments)  
18.575 Million cell updates/sec

Title: US-09-724-406-26

Perfect score: 578

Sequence: 1 DIVLTQSPASLAVSLGRAT.....CQHSGLPFTFGSGTKLEIK 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_0601.\*
- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*
  - 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*
  - 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*
  - 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*
  - 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*
  - 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*
  - 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*
  - 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*
  - 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*
  - 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*
  - 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*
  - 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*
  - 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*
  - 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*
  - 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*
  - 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*
  - 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*
  - 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*
  - 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*
  - 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.\*
  - 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*
  - 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	526	91.0	212	15	AA052659
2	524	90.7	111	21	AA010019
3	521	90.1	131	13	AA028668
4	520	90.0	111	17	AA085236
5	513.5	88.8	110	15	AA060564
6	513.5	88.8	110	20	AA039532
7	512	88.6	112	16	AA080272
8	510	88.2	111	22	AA069661
9	509	88.1	125	12	AA015440
10	508	87.9	111	15	AA060528
11	505	87.4	112	16	AA075431

12	502.5	86.9	151	22	AA035100	Antibody variable
13	502	86.9	132	17	AA003723	Anti-human gp39 MA
14	500	86.5	218	16	AA075457	Mouse antibody PB3
15	500	86.5	218	16	AA075459	Mouse antibody F4-
16	494	85.5	437	19	AA037738	Nucleotide sequenc
17	489.5	84.7	146	18	AA019578	Mouse anti-idiotyp
18	486.5	84.2	124	12	AA015442	Light chain variab
19	485	83.9	111	10	AA090541	Immunoglobulin L c
20	485	83.9	131	10	AA090543	Amino acids sequen
21	484.5	83.8	110	21	AA090821	317G5 hybridoma VL
22	484.5	83.8	112	11	AA007456	N-terminal sequenc
23	482.5	83.5	110	18	AA022952	Monoclonal antibod
24	482.5	83.5	110	20	AA086119	Murine 340 VI amin
25	480	83.3	111	15	AA055123	Mouse anti-HIV mu5
26	480	83.3	111	15	AA060302	Anti HIV antibody
27	480	83.3	238	19	AA083042	Anti-Fas MAB HFE7A
28	480	83.3	238	21	AA014748	Mouse anti-Fas ant
29	480	83.3	238	21	AA090898	Murine anti-Fas an
30	479	82.9	111	21	AA059267	Antibody 4H5 L cha
31	479	82.9	111	21	AA051144	Murine derived pro
32	479	82.9	111	21	AA051146	Murine derived pro
33	479	82.9	132	16	AA070189	Mouse MAB 389 ligh
34	479	82.9	132	20	AA023767	Light chain variab
35	479	82.9	132	20	AA018120	Light chain sequen
36	479	82.9	305	21	AA059264	Antibody 4H5 H cha
37	479	82.9	305	21	AA059265	Antibody 4H5 L cha
38	479	82.9	305	21	AA051141	Murine derived pro
39	479	82.9	305	21	AA051142	Murine derived pro
40	476.5	82.4	110	16	AA074957	Immunoglobulin lig
41	476.5	82.4	130	16	AA074957	Anti-idiotyp anti
42	473.5	81.9	110	16	AA074956	Immunoglobulin lig
43	473.5	81.9	146	16	AA074966	Anti-idiotyp anti
44	473	81.8	113	22	AA071895	Monoclonal antibod
45	472	81.7	131	18	AA034516	Variable kappa cha

## ALIGNMENTS

RESULT	1
AA052659	AA052659 standard; Protein; 212 AA.
ID	AA052659 standard; Protein; 212 AA.
XX	
AC	AA052659;
XX	
DT	16-JAN-1995 (first entry)
XX	
DE	Porphyryn antibody light chain.
XX	
KW	Light; heavy; chain; monoclonal antibody; porphyrin;
KW	absorption spectrum; oxidase activity.
XX	
OS	Mus musculus.
XX	
PN	JP06105686-A.
XX	
PD	19-APR-1994.
XX	
PF	28-SEP-1992; 92JP-0258584.
XX	
PR	28-SEP-1992; 92JP-0258584.
XX	
PA	(TOYM ) TOYOBO KK.
XX	
DR	WPI; 1994-163128/20.
XX	N-PSDB; AA062631.
PT	Gene encoding monoclonal antibody against porphyrin - useful to
PT	produce antibody with modified properties, to interact with e.g.
PT	P450, haemoglobin; etc.
XX	
PS	Claim 4; Page 10-11; 16pp; Japanese.
XX	

CC The sequences given in AAR52659-60 represent the light and heavy chains  
 CC respectively of a monoclonal antibody which specifically recognises  
 CC a specified porphyrin molecule. Using the genes encoding these  
 CC proteins, the porphyrin antibody can be mass produced inexpensively.  
 CC Features of the porphyrin monoclonal antibody may be modified, eg.  
 CC absorption spectrum shift and expression of oxidase activity, by  
 CC alterations within the coding region.

XX Sequence 212 AA;  
 SQ Query Match 91.0%; Score 526; DB 15; Length 212;  
 Best Local Similarity 91.0%; Pred. No. 1.1e-37;  
 Matches 101; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNLDES 60  
 Db 1 divltqspaslavslgqratiscrasksvasgylymhwqgpgppkllislatnles 60  
 QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111  
 Db 61 gvparfsgsggtdftlnhpveeedvatyycqhsrelptfgagtklelk 111

RESULT 2  
 AAB10019  
 ID AAB10019 standard; Protein; 111 AA.

AC AAB10019;  
 XX 01-NOV-2000 (first entry)  
 DT H. pylori HSP60-binding antibody light chain protein.  
 DE Acid-resistant microorganism; detection; faecal; intestine; infection;  
 KW monoclonal antibody; heat shock protein; HSP60; light chain.

XX Unidentified.  
 XX WO200026671-A1.  
 XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-EP09212.  
 XX 29-OCT-1998; 98EP-0120517.  
 PR 06-NOV-1998; 98EP-0120687.  
 XX (CONN-) CONNEX GMBH.  
 XX Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;  
 PI Ringels A;  
 XX WPI; 2000-365747/31.  
 DR N-PSDB; AAA40199.

XX Detecting infection by acid-fast microbes for diagnosis of Helicobacter  
 PT pylori, comprises reacting a faecal sample with two binding reagents for  
 PT antigens that survive intestinal passage

XX Disclosure; Fig 1; 84pp; German.

XX This invention describes a novel method for the detection of a mammalian  
 CC infection by an acid-resistant microorganism (A) by treating a faecal  
 CC sample with at least two different monoclonal antibodies (Mab) (or their  
 CC fragments or derivatives) or aptamers (collectively (I)) and detecting  
 CC formation of a complex (C) between (I) and the corresponding antigen of  
 CC (A). The first and second (I) bind to epitopes of different antigens  
 CC (Ag). These epitopes are present, after passage through the intestines,  
 CC in at least some mammals, and have either: (i) their native structure;  
 CC or (ii) a structure against which an antibody is produced by an animal  
 CC infected or immunized with (A), or its extract, lysate, derived protein  
 CC or fragment, or with a synthetic peptide. Practically all mammals display

CC at least one of the specified epitopes. The method is used to detect  
 CC infection by acid-fast bacteria, particularly of the genera Helicobacter,  
 CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,  
 CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used  
 CC therapeutically. The method is direct and non-invasive, and provides an  
 CC inexpensive and easily standardizable diagnosis, despite possible  
 CC degradation of antigens during passage through the intestines. This  
 CC sequence represents the H. pylori heat shock protein, HSP60-binding  
 CC antibody (DMS ACC2356) light chain which is used to illustrate the method  
 CC of the invention.

XX Sequence 111 AA;

Query Match 90.7%; Score 524; DB 21; Length 111;  
 Best Local Similarity 90.1%; Pred. No. 8.4e-38;  
 Matches 100; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNLDES 60  
 Db 1 divltqspaslavslgqratiscrasksvstsgysyihwyqgpgppklliflasnles 60  
 QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111  
 Db 61 gvparfsgsggtdftlnhpveeedaatyhcqhsrelptfgagtklelk 111

RESULT 3  
 AAR28668  
 ID AAR28668 standard; Protein; 131 AA.

AC AAR28668;  
 XX 30-MAR-1993 (first entry)  
 DT p12-k2.

XX Human; antibody; interleukin-6; receptor; IL-6R; light chain; L; H;  
 KW heavy chain; variable region; mouse; monoclonal; hybridoma; AUK12-20;  
 KW plasmid; p12-k2; p12-h2.

XX Synthetic.

XX Key Location/Qualifiers  
 FH Peptide 1..20  
 FT /note= "Signal peptide"  
 FT Protein 21..131  
 FT /note= "Mature peptide"

XX WO9219759-A.  
 XX 12-NOV-1992.  
 XX 24-APR-1992; 92WO-JP00544.  
 XX 25-APR-1991; 91JP-0095476.  
 PR 19-FEB-1992; 92JP-0032084.  
 XX (CHUS) CHUGAI SEIYAKU KK.

XX Bendig MM, Jones ST, Saidanha JW, Sato K, Tsuchiya M;

XX WPI; 1992-398882/48.  
 DR N-PSDB; AAQ30753.

XX Reconstituted human antibody to human interleukin-6 receptor -  
 PT has low antigenicity and contains mouse V-region complementarity  
 PT determining regions

XX Disclosure; Page 118; 207pp; Japanese.

XX The sequences given in AAR28668-69 were encoded by plasmid sequences  
 CC which were used in example to illustrate the production of a human

CC antibody which recognises human interleukin-6 receptor (IL-6R). The  
 CC antibody comprises light (L) chain and heavy (H) chain variable  
 CC regions which were derived from a mouse monoclonal antibody produced  
 CC from the hybridoma AUK12-20 which contained the plasmids p12-k2 and  
 CC p12-h2.

XX Sequence 131 AA;

Query Match 90.18; Score 521; DB 13; Length 131;  
 Best Local Similarity 90.18; Pred. No. 1.8e-37;  
 Matches 100; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Oy 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQOKAGOPPKLLIHLASNL 60  
 Db 21 divltqspaslgvslgratiscrasksvstsgysymhwyyqgkpgqpkllilylasnles 80

Oy 61 GVPARFSGSGGTDFTLNIHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
 Db 81 gvparfsgsgsgtdftlnihpveedaatyycqhsrenpytfgggtkleik 131

#### RESULT 4

AAR85236  
 ID AAR85236 standard; Protein; 111 AA.

XX AAR85236;

DT 05-APR-1996 (first entry)

DE Murine 206 antibody variable light chain region.

KW Antibody; varicella zoster virus; Ab; VSV; treatment; prevention;  
 KW infection; detection; isolation; monoclonal antibody; MAB; mimetic;  
 KW humanised; murine; heavy chain; light chain; variable region; CDR;  
 KW NEMW; REI.

XX Mus musculus.

FH Key Location/Qualifiers

FT Binding-site 24..38

FT Binding-site /label= CDR 1.

FT Binding-site 54..60

FT Binding-site /label= CDR 2.

FT Binding-site 93..101

FT Binding-site /label= CDR 3.

XX WO9531546-A1.

XX 23-NOV-1995.

XX 28-APR-1995; 95WO-US05305.

XX 28-APR-1994; 94US-0234133.

XX (SCOT-) SCOTGEN BIOPHARMACEUTICALS INC.

XX Carr FJ, Harris WJ, Moss MT, Wallace TP;

XX WPI; 1996-010932/01.

XX N-PSDB; AAT06051.

PT Varicella zoster virus gpIII antibody with humanised variable region  
 PT - for treatment, prevention or diagnosis of varicella zoster virus  
 PT infections

XX Claim 2; Figure 2; 58pp; English.

XX Antibodies (Ab) comprising a humanised variable region specifically  
 CC binding to a varicella zoster virus (VZV) can be used to treat or  
 CC prevent VZV infection. They may optionally be attached to a  
 CC therapeutic agent. They may also be used, when labelled, to detect  
 CC VZV particles and infected cells in blood; for the removal/  
 CC

CC neutralisation of infectious VZV in blood; for the selection/  
 CC isolation of human monoclonal Abs and for the design of (non-)  
 CC peptide mimetics with similar diagnostic and therapeutic uses. The  
 CC variable region CDR portion is derived from murine Ab 206 and the  
 CC heavy and light chain variable region frameworks are NEMW and REI  
 CC respectively. The framework may include at least one mutation that  
 CC improves binding specificity or affinity. The heavy and light chain  
 CC variable regions are preferably (V2VHu-VH, -VHKFIIS, -VHTLY,  
 CC -VHKVRSE or -VHS) and V2VHuVH respectively.

XX Sequence 111 AA;

Query Match 90.0%; Score 520; DB 17; Length 111;  
 Best Local Similarity 91.0%; Pred. No. 1.8e-37;  
 Matches 101; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Oy 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQOKAGOPPKLLIHLASNL 60  
 Db 1 divltqspas..avslgratiscrasksvstsgysymhwyyqgkpgqpkllilylasnles 60

Oy 61 GVPARFSGSGGTDFTLNIHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
 Db 61 gvlsrfsgsgsgtdftlnihpveedaatyycqhsrelptfgpgtkleik 111

#### RESULT 5

AAR60564

ID AAR60564 standard; Protein; 110 AA.

XX AAR60564;

DT 25-APR-1995 (first entry)

DE Anti-carcinoembryonic antigen chimeric light chain Ab.

XX Anti-carcinoembryonic antigen chimeric light chain Ab.  
 KW Anti-carcinoembryonic antigen chimeric antibodies; CEAS;  
 KW chimeric human-murine; breast or colorectal carcinoma;  
 KW light chain.

XX Chimeric Mus muscaris.

OS Chimeric Homo sapiens.

XX WO9419466-A.

XX 01-SEP-1994.

XX 16-FEB-1994; 94WO-US01709.

XX 16-FEB-1993; 93US-0017570.

XX (DOWC ) DOW CHEM CO.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Gourlie BB, Kaplan DA, Mezes PS, Rixon MW, Schlom J;

XX WPI; 1994-294331/36.

XX N-PSDB; AAQ71394.

PT Anti-carcinoembryonic antigen chimeric antibodies - for diagnosis  
 PT and therapy of carcinoma, e.g. breast or colorectal carcinoma  
 XX Claim 11; Page 50; 67pp; English.

XX AAQ71394 codes for AAR60564 the antibody light chain region of  
 CC murine-human anti-carcinoembryonic antigen (CEA) chimeric  
 CC antibody. Which can be used in in vitro immunoassays for  
 CC the detection of CEA, and monitoring of tumour-associated  
 CC antigen during therapy. It can also be used in vivo diagnostically,  
 CC or in therapy for the treatment of tumours associated with  
 CC colorectal and breast carcinomas, as well those of the  
 CC gastrointestinal tract, lung, ovary and pancreas.

SQ Sequence 110 AA;

Query Match 88.8%; Score 513.5; DB 15; Length 110;

Best Local Similarity 90.1%; Pred. No. 6.5e-37;

Matches 100; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60

DB 1 divltspasltvslglratiscrasksvasgysymhwyrpqpqppklliyasnlqs 60

QY 61 GVPARFSGSGGDTFTLNHPVEEEDASTYCYQHSGELPFTFGSGTKLEIK 111

DB 61 gvparfsgsgsgdtftlnhpveeedaatyqchsrelp-tfgggtkleik 110

RESULT 6

AAV39532

ID AAV39532 standard; protein; 110 AA.

XX

AC AAV39532;

XX

DT 22-NOV-1999 (first entry)

XX

DE Murine COL1 VK chain (COL1MuVK).

XX

KW Antibody; human; carcinoembryonic antigen; CEA; cancer; tumour imaging;

KW tumour identification; metastasis; diagnosis; mouse; VK chain.

XX

OS Mus sp.

XX

PN WO9943817-A1.

XX

PD 02-SEP-1999.

XX

PF 25-FEB-1998; 98WO-US03680.

XX

PR 25-FEB-1998; 98WO-US03680.

XX

PA (DOWC ) DOW CHEM CO.

XX

PI Anderson WHK, Tempest PR, Carr FJ, Harris WJ, Armour K;

XX

DR WPI; 1999-550870/46.

XX

PT New antibody specific for treatment and diagnosis of cancer

XX

PS Claim 2; Fig 2; 82pp; English.

XX

CC This sequence represents the murine COL1 VK chain (COL1MuVK), and was

CC used to obtain the humanised antibody (Ab) of the invention. The

CC humanised Ab, or its fragment, specifically binds to carcinoembryonic

CC antigen (CEA). The Ab, optionally coupled to an effector or label, is

CC used to treat or prevent CEA-expressing cancers (e.g. of breast, ovary,

CC lung, stomach, or colon) and to detect CEA-expressing cells, either in

CC vitro (optionally on a solid support) or in vivo (particularly by tumour

CC imaging to identify tumours and metastases before surgery), for diagnosis

CC or prognosis. It is not significantly immunogenic, i.e. the Ab does not

CC induce a human anti-murine antibody or allergic response, or non-specific

CC cytotoxicity, so can be administered repeatedly. It retains specificity

CC for CEA, and has improved clearance (allowing efficient targeting) and

CC metabolic properties.

XX

SQ Sequence 110 AA;

Query Match

Best Local Similarity 90.1%; Pred. No. 6.5e-37;

Matches 100; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60

DB 1 divltspasltvslglratiscrasksvasgysymhwyrpqpqppklliyasnlqs 60

QY 61 GVPARFSGSGGDTFTLNHPVEEEDASTYCYQHSGELPFTFGSGTKLEIK 111

DB 61 gvparfsgsgsgdtftlnhpveeedaatyqchsrelp-tfgggtkleik 110

RESULT 7

AAV39532

ID AAV39532 standard; protein; 112 AA.

XX

AC AAV39532;

XX

DT 27-FEB-1996 (first entry)

XX

DE V1 region from an antibody against cancer-specific mucin.

XX

KW Antibody; heavy chain; light chain; variable region; cancer; mucin;

KW hybridoma cell; murine; mouse; pancreatic cancer cell; expression vector;

KW Fv; human; constant domain; chimera; anaphylaxis.

XX

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Domain 24..38

FT /label= CDR1

FT Domain 54..60

FT /label= CDR2

FT Misc-difference 64

FT /note= "encoded by GAG"

FT Domain 93..101

FT /label= CDR3

XX

PN FR2714915-A1.

XX

PD 13-JUL-1995.

XX

PF 13-JAN-1995; 95FR-0000349.

XX

PR 13-JAN-1994; 94JP-0002131.

XX

PA (TOYJ ) TOSOH CORP.

XX

PI Chung Y, Iba Y, Kaneko T, Sowa M, Yasukawa K;

XX

DR WPI; 1995-247908/33.

XX

DR N-PSDB; AAQ98534.

XX

PT New variable domains of antibody recognising cancer specific mucin

PT - and related DNA and expression vectors, producing chimeric

PT mouse-human antibody for diagnosis and treatment of cancer

XX

PS Claim 6; Page 16-17; 25pp; French.

XX

CC The amino acid sequence of the variable region from the light chain of

CC an antibody against cancer-specific mucin. The coding sequence was

CC isolated from N2D hybridoma cells expressing a murine antibody reactive

CC with pancreatic cancer cells. The DNA encoding the antibody variable

CC regions from the heavy (AAQ98533) and light chains were isolated and

CC inserted into vectors. These vectors express the domains as an Fv

CC antibody. Vectors which additionally contain genes encoding the human

CC constant domains express a chimeric mouse-human antibody. The antibodies

CC are useful in the detection and treatment of cancer. The chimeras should

CC be less likely to cause anaphylaxis than the original murine antibody.

XX

SQ Sequence 112 AA;

Query Match

Best Local Similarity 88.6%; Score 512; DB 16; Length 112;

Matches 97; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60

DB 1 divltspasltvslglratiscrasksvasgysymhwyrpqpqppklliyasnlqs 60

Db 1 divltqspaslavslgqratiscrasksvtsdtsfymhwyyqkpgqppklllylasnles 60  
QY 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
DB 61 gvpdrfsgsggtdftlnhpveedaatyqcshsrefpwtfggggkcleik 111

RESULT 8  
AAB69661  
ID AAB69661 standard; Protein; 111 AA.  
XX  
AC AAB69661;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE Murine Fd79 antibody light chain SEQ ID NO: 46.  
XX  
KW Humanised immunoglobulin; mouse; human; antibody; heavy chain; diabetes;  
KW light chain; graft versus host disease; transplant; autoimmune disease;  
KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
KW myasthenia gravis; herpes infection; myeloid leukaemia; CMV infection.  
XX  
OS Mus sp.  
XX  
PN US6180370-B1.  
XX  
PD 30-JAN-2001.  
XX  
PF 07-JUN-1995; 95US-0484537.  
XX  
PR 28-DEC-1988; 88US-0290975.  
PR 13-FEB-1989; 89US-0310252.  
PR 28-SEP-1990; 90US-0590274.  
PR 19-DEC-1990; 90US-0634278.  
XX  
PA (PROT-) PROTEIN DESIGN LABS INC.  
XX  
PI Queen CL, Sellick HE;  
XX  
DR WPI; 2001-190856/19.  
XX  
PT Producing humanized immunoglobulin, involves producing a cell  
PT containing DNA segments encoding humanized heavy and light chain  
PT variable regions, and expressing the DNA segments in the cell -  
XX  
PS Example 6; Fig 2; 145pp; English.  
XX  
CC The present invention describes a method of producing humanised  
CC immunoglobulins involving expressing in a cell a nucleic acid encoding a  
CC humanised version of an immunoglobulin. This is obtained by comparing a  
CC donor and human immunoglobulin and producing a combined antibody which  
CC contains part of each. These are useful in the treatment of  
CC graft-versus-host disease, transplant rejection, autoimmune diseases such  
CC as diabetes, rheumatoid arthritis, myasthenia gravis, multiple sclerosis  
CC and systemic lupus erythematosus, herpes infections, CMV virus infections  
CC and myeloid leukaemia. The present sequence is an antibody used to  
CC demonstrate the method of the invention.  
XX  
SQ Sequence 111 AA;

Query Match 88.2%; Score 510; DB 22; Length 111;  
Best Local Similarity 88.3%; Pred. No. 1.3e-36;  
Matches 98; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQQAGQPPKLLIHLASNLES 60  
DB 1 divltqspaslavslgqratiscrasksvtsdtsfymhwyyqkpgqppklllylasnles 60

QY 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
DB 61 gvpdrfsgsggtdftlnhpveedaatyqcshsweipwtfggggkcleik 111

RESULT 9  
AAR15440  
ID AAR15440 standard; Protein; 125 AA.  
XX  
AC AAR15440;  
XX  
DT 25-FEB-1992 (first entry)  
XX  
DE Light chain variable region of MAb 2C2.  
XX  
KW HRV; ICAM-1; antigen-binding fragment; inflammation;  
KW auto-immune disease.  
XX  
OS Homo sapiens.  
XX  
PN EP459577-A.  
XX  
PD 04-DEC-1991.  
XX  
PF 25-MAY-1991; 91EP-0201243.  
XX  
PR 01-JUN-1990; 90US-0532001.  
XX  
PA (MERI) MERCK & CO INC.  
XX  
PI Colonna RJ, Condra JH, Tomassini JE, Sardana VV;  
XX  
DR WPI; 1991-355850/49.  
XX  
PT Microbially expressed portions of monoclonal antibody - can block  
PT attachment of rhinovirus ligands to inter-cellular adhesion  
PT molecule (ICAM-1)  
XX  
PS Claim 1; Page 20; 28pp; English.  
XX  
CC This is one of six antibody fragments from MAb's specific for domain  
CC 1 of ICAM-1. The peptide fragments can be used to treat or prevent  
CC rhinovirus infection. See AAR15437-R15443.  
XX  
SQ Sequence 125 AA;

Query Match 88.1%; Score 509; DB 12; Length 125;  
Best Local Similarity 87.4%; Pred. No. 1.8e-36;  
Matches 97; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQQAGQPPKLLIHLASNLES 60  
DB 1 divltqspaslavslgqratiscrasksvtsdtsfymhwyyqkpgqppklllylasnles 60

QY 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
DB 61 gvpdrfsgsggtdftlnhpveedaatyqcshsweipwtfggggkcleik 111

RESULT 10  
AAR60528  
ID AAR60528 standard; Protein; 111 AA.  
XX  
AC AAR60528;  
XX  
DT 07-APR-1995 (first entry)  
XX  
DE Light chain variable region of monoclonal antibody 5E8.  
XX  
KW Secretion sequence; vector; antibody; single chain antibody;  
KW promoter; Bacillus subtilis; ss.  
XX  
OS Synthetic.  
XX  
PN EP614982-A.  
XX

PD	14-SEP-1994.	
XX		
PF	05-MAR-1994; 94EP-0200564.	
XX		
XX	10-MAR-1993; 93IT-OMI0456.	
XX	(ENIE ) ENIRICERCH SPA.	
XX		
PI	Cuzzoni A, De FERRA F, Grandi G, Tortora O, Tosi C;	
DR	WPI; 1994-281209/35.	
DR	N-PSDB; AAQ71286.	
XX		
PT	Vector for secretion of single chain antibody from Bacillus	
PT	subtilis - confg. neutral protease promoter and specific leader	
PT	sequence, providing high yield of soluble antibody for diagnostic	
PT	or therapeutic use	
XX		
PS	Example 3; Figure 2; 27pp; English.	
XX		
CC	The sequence encoding the variable region of the light chain of the	
CC	monoclonal antibody 5E8 (specific for the alpha subunit of human	
CC	gonadotropin) was amplified for its use in a recombinant vector	
CC	which also comprised (1) the promoter of the gene for the neutral	
CC	protease of Bacillus subtilis BGSC 1A341, (2) the secretion sequence	
CC	described in AAQ71278. The amplified fragment was used to produce	
CC	DNA sequence encoding an antibody in single molecule form with the	
CC	sequence VH/VK-L-VK/VH-(TAG)n, where VH and VK are the variable	
CC	regions of the heavy and light chains of the antibody and L is the	
CC	linker of sequence Val-Ser-Ser-(Gly(4)-Ser)3. TAG is a peptide	
CC	recognised by polyclonal antibodies directed towards the same peptide	
CC	n is 1 or 0. The vector is used for the production of single chain	
CC	antibodies which have improved pharmacokinetic properties and can	
CC	produced more economically than monoclonal antibodies. The vector	
CC	allows production of antibodies in completely soluble form with	
CC	secretion in high yield.	
XX		
SQ	Sequence 111 AA;	
XX		

```

Query Match      87.9%; Score 508; DB 15; Length 111;
Best Local Similarity 89.1%; Pred. No. 1.9e-36;
Matches 98; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY   1 DIVLTQSPLAVSLGQRATISCRASKSVASGYNYMHWIQKAGOPPKLIHLASNLES 60
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1 diqqqtspaslvslgqraticrasksvasgyvnywqkpgppklilylasnles 60

QY   61 GVPARFSSGGSGTDFTLNIHPEVEDASTYYVCQHSGLPFTFGSGTKLEI 110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   61 gvparfsgsgsgtdftlnhpveedaatyvcghsrelpitfgagtklei 110

RESULT 11
AAR75431 AAR75431 standard; Protein; 112 AA.
XX XX
AC AC
DT DT
XX XX
DE DE
XX XX
KW KW
XX XX
OS OS
XX XX
Key Location/Qualifiers
FH Region 24...38
FT /note= "Complementarity determining region (CDR) 1"
FT Region 54...60
FT /note= "CDR2"

```

```

Region          93..101
                /note= "CDR3"
FT             XX
FT             XX
FT             XX
FT             XX
FT             XX
PD             XX
PD             XX
PF             XX
XX             28-JUN-1993;   93JP-0156707.
XX             28-JUN-1993;   93JP-0156707.
XX             (TOYJ ) TOSOH CORP.
XX             PA
XX             XX
DR             WPI; 1995-220118/29.
DR             N-PSDB; AAQ/4148.
XX             XX
PT             DNA encoding an antibody recognising human thyroid-stimulating
PT             hormone - and preparation of chimeric antibody by expressing the
PT             DNA in a transformed host cell
PT             XX
XX             Claim 8; Fig 4; 8pp; Japanese.
XX             XX
CC             The sequence is the light chain variable region of an antibody
CC             recognising human thyroid stimulating hormone (hTSH). Prepn. of a
CC             chimeric antibody by expressing the DNA in a transformed host cell
CC             is also claimed. The anti-hTSH antibody has an additional useful
CC             function which could not be given by a mouse derived anti-TSH monoclonal
CC             antibody (sic).
XX             XX
SQ             Sequence      112 AA;

Query Match           87.4%; Score 505; DB 16; Length 112;
Best Local Similarity 86.3%; Pred. No. 3.5e-36;
Matches 96; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY    1 DIVLTQPASLAVSLGORATISCRASKSVSAGSYNYMHWTQQKAGPPKLLIHLASNLES 60
      |||||||
Db     1 divltqpaslpvslqraticscragssvsyssymhwyyqkpgqpklifikfasnles 60
      |||||||

QY    61 GVPARFSGSGCGDFTLNHPVEEDASTYYCQHSGELPFTFGSGTKLEIK 111
      |||||||
Db     61 gvparfsgsgdftlnhpveeddaatyycqtweiprtfgggtkleik 111

RESULT 12
AAB35100
ID    AAB35100 standard; Protein; 151 AA.
XX
AC    AAB35100;
XX
DT    27-MAR-2001 (first entry)
XX
DE    Antibody variable region fusion protein #4.
XX
KW    Filamentous phage; protein display; pVII; pIX;
KW    combinatorial antibody library.
XX
OS    Synthetic.
XX
PN    WO200071694-A1.
XX
PD    30-NOV-2000.
XX
XX    24-MAY-2000; 2000WO-US14433.
XX
XX    25-MAY-1999; 99US-0318786.
XX
PA    (SCRI ) SCRIPPS RES INST.
XX
PI    Janda KD, Wirsching P, Lerner RA, Gao C;
XX
DR    WPI; 2001-032030/04.
```





```

PI Barsomian G, Copeland DP, Hillhouse D, Johnson T;
XX WPI; 1995-224291/29.
DR N-PSDB; AAQ92501.
XX
XX Generating new antibodies specific for immunorecessive epitopes -
PT by selection from variegated V gene library cloned from
PT immuno:tolerance derived antibody repertoire, useful in diagnosis,
PT purifcn. and therapy, e.g. of cancer
XX
PS Disclosure; Page 76-77; 109pp; English.
XX
XX The sequence of the light chain variable region from the mouse antibody
CC FB3-2. This sequence was isolated from a variegated display library
CC (VDL) of variable regions derived from a repertoire of antibodies from
CC an immunotolerised animal. The VDL is generated by PCR amplifying the
CC variable regions from the antibody coding sequences using the primers
CC AAQ74153-74. The variable regions, esp the complementarity determining
CC regions (CDR; see AAR75462-93 for examples of CDRs) from the
CC immunotolerant animals' antibodies are used to construct an antibody
CC against a immunorecessive antigen e.g. a cell surface marker on a foetal,
CC cancer or stem cell, which can differentiate between variant or related
CC forms of the antigen. The antibodies generated can be used in the
CC diagnosis, e.g. detection of the immunorecessive antigen, or in therapy
CC e.g. of cancer, Alzheimer's disease or familial hypercholesterolaemia.
CC The method of production of the antibody allows rapid and sensitive
CC isolation of antibodies that would be difficult to isolate by standard
CC methods. The antibodies produced have greater binding affinity than
CC those produced by combinatorial/hybridoma methods.
XX
XX Sequence 218 AA;

Query Match 86.5%; Score 500; DB 16; Length 218;
Best Local Similarity 84.7%; Pred. No. 1.9e-35;
Matches 94; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVSASGYNYMHVYQKAGOPPKLLIHLASNLES 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
AAR75459
ID AAR75459 standard; Protein; 218 AA.
XX
AC AAR75459;
XX
XX 07-FEB-1996 (first entry)
XX
DE Mouse antibody F4-7 light chain variable region protein sequence.
XX
KW Primer; amplification; PCR; mouse; kappa chain; heavy chain; Fab;
KW antibody; immunotolerance; animal; variegated display library;
KW variable region; antigen; immunorecessive; cell surface marker; foetal;
KW cancer; stem cell; variant; therapy; Alzheimer's disease; hybridoma;
KW familial hypercholesterolaemia; binding affinity.
XX
OS Mus musculus.
XX
PN W09515982-A2.
XX
PD 15-JUN-1995.
XX
PF 08-DEC-1994; 94WO-US14106.
XX
PR 06-DEC-1994; 94US-0350400.
PR 08-DEC-1993; 93US-0164022.
XX

```

```

PA (GENZ ) GENZYME CORP.
XX Barsomian G, Copeland DP, Hillhouse D, Johnson T;
XX WPI; 1995-224291/29.
DR N-PSDB; AAQ92503.
XX
XX Generating new antibodies specific for immunorecessive epitopes -
PT by selection from variegated V gene library cloned from
PT immuno:tolerance derived antibody repertoire, useful in diagnosis,
PT purifcn. and therapy, e.g. of cancer
XX
PS Disclosure; Page 80-81; 109pp; English.
XX
XX The sequence of the light chain variable region from the mouse antibody
CC F4-7. This sequence was isolated from a variegated display library (VDL)
CC of variable regions derived from a repertoire of antibodies from an
CC immunotolerised animal. The VDL is generated by PCR amplifying the
CC variable regions from the antibody coding sequences using the primers
CC AAQ74153-74. The variable regions, esp the complementarity determining
CC regions (CDR; see AAR75462-93 for examples of CDRs) from the
CC immunotolerant animals' antibodies are used to construct an antibody
CC against a immunorecessive antigen e.g. a cell surface marker on a foetal,
CC cancer or stem cell, which can differentiate between variant or related
CC forms of the antigen. The antibodies generated can be used in the
CC diagnosis, e.g. detection of the immunorecessive antigen, or in therapy
CC e.g. of cancer, Alzheimer's disease or familial hypercholesterolaemia.
CC The method of production of the antibody allows rapid and sensitive
CC isolation of antibodies that would be difficult to isolate by standard
CC methods. The antibodies produced have greater binding affinity than
CC those produced by combinatorial/hybridoma methods.
XX
XX Sequence 218 AA;

```

```

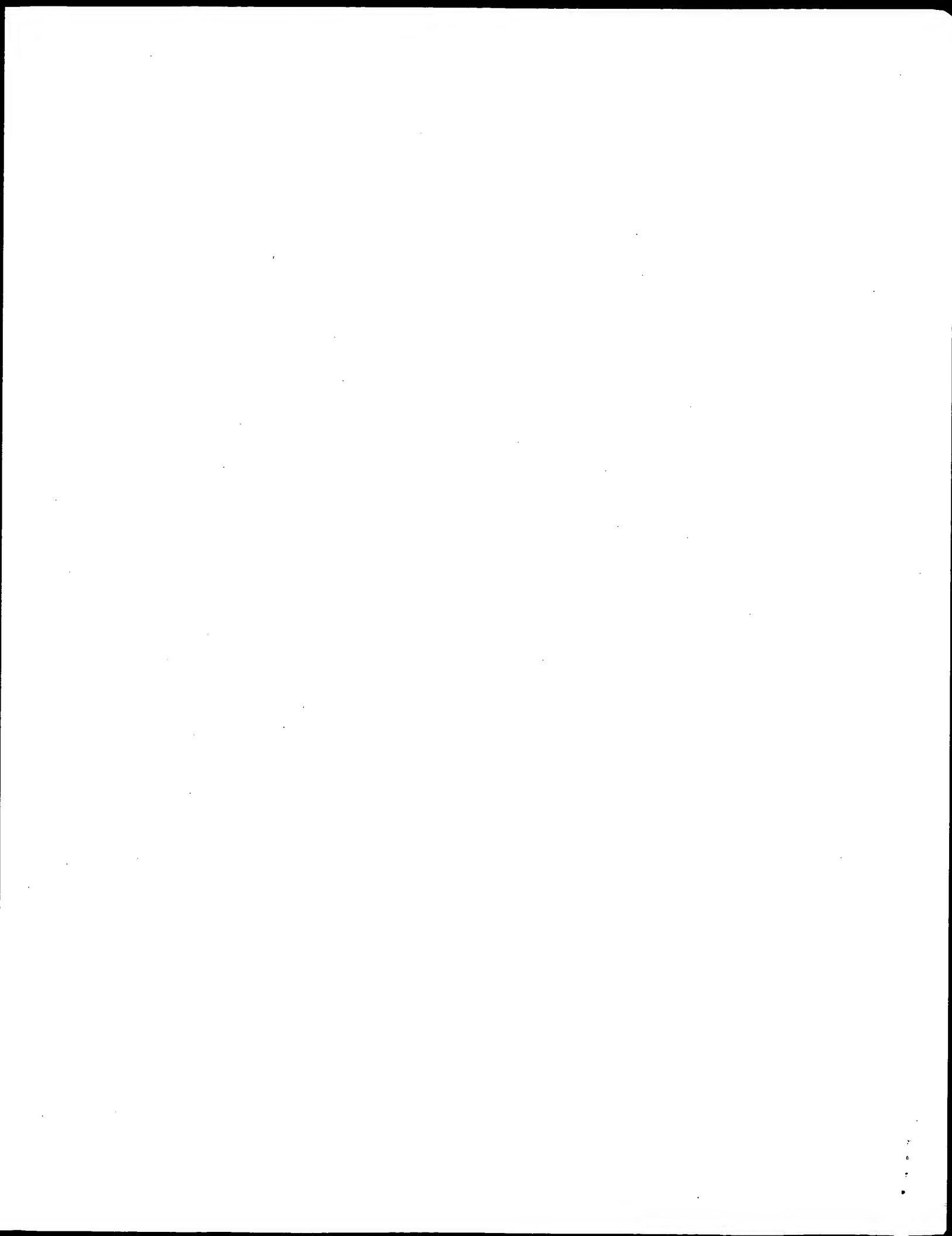
Query Match 86.5%; Score 500; DB 16; Length 218;
Best Local Similarity 84.7%; Pred. No. 1.9e-35;
Matches 94; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVSASGYNYMHVYQKAGOPPKLLIHLASNLES 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: June 28, 2001, 16:14:38  
Job time: 1326 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:01:15 ; Search time 138.34 Seconds  
(without alignments)  
16.163 Million cell updates/sec

Title: US-09-724-406-26  
Perfect score: 578  
Sequence: 1 DIVLTQSPASVSLQRAT.....CQHSGLPPTFGSGTKLEIK 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	521	90.1	131	1	US-08-137-117D-25
2	521	90.1	131	2	US-08-436-717-25
3	513.5	88.8	110	1	US-08-017-570-2
4	513.5	88.8	110	1	US-08-471-426-2
5	513.5	88.8	110	5	PCT-US94-01709-2
6	510	88.2	111	1	US-07-634-278-46
7	510	88.2	111	1	US-08-477-728-46
8	510	88.2	111	1	US-08-474-040-46
9	510	88.2	111	1	US-08-487-200-46
10	510	88.2	111	4	US-08-484-537-46
11	508	87.9	111	1	US-08-207-169A-4
12	505	87.4	111	5	PCT-US94-14106-53
13	502	86.9	132	2	US-08-379-057-16
14	500	86.5	218	5	PCT-US94-14106-57
15	489.5	84.7	146	2	US-08-653-402B-12
16	484.5	83.8	110	3	US-08-483-749A-20
17	479	82.9	132	2	US-08-483-636-2
18	479	82.9	132	2	US-08-483-632-2
19	475	82.2	111	1	US-08-491-845-8
20	474	82.0	131	4	US-08-579-378A-14
21	472	81.7	131	3	US-08-589-939-3
22	461	79.8	106	3	US-08-466-151-6
23	460.5	79.7	110	1	US-08-442-542-8
24	460.5	79.7	110	1	US-08-765-469-8
25	453	78.4	120	1	US-08-111-080-26
26	453	78.4	120	1	US-08-211-980-26
27	453	78.4	120	5	PCT-US93-07967-26

28 452.5 78.3 112 2 US-08-888-366-18 Sequence 18, Appl  
29 450 77.9 111 1 US-08-275-053-11 Sequence 11, Appl  
30 450 77.9 111 3 US-08-466-151-2 Sequence 2, Appl  
31 450 77.9 121 1 US-08-111-080-22 Sequence 22, Appl  
32 450 77.9 121 1 US-08-211-980-22 Sequence 22, Appl  
33 450 77.9 121 5 PCT-US93-07967-22 Sequence 22, Appl  
34 447 77.3 120 1 US-08-111-080-24 Sequence 24, Appl  
35 447 77.3 120 1 US-08-211-980-24 Sequence 24, Appl  
36 447 77.3 120 5 PCT-US93-07967-24 Sequence 24, Appl  
37 446 77.2 111 1 US-07-634-278-54 Sequence 54, Appl  
38 446 77.2 111 1 US-08-477-728-54 Sequence 54, Appl  
39 446 77.2 111 1 US-08-474-040-54 Sequence 54, Appl  
40 446 77.2 111 1 US-08-487-200-54 Sequence 54, Appl  
41 446 77.2 111 4 US-08-484-537-54 Sequence 54, Appl  
42 446 77.2 131 1 US-07-634-278-67 Sequence 67, Appl  
43 446 77.2 131 1 US-08-477-728-67 Sequence 67, Appl  
44 446 77.2 131 1 US-08-474-040-67 Sequence 67, Appl  
45 446 77.2 131 1 US-08-487-200-67 Sequence 67, Appl

#### ALIGNMENTS

RESULT 1  
US-08-137-117D-25  
; Sequence 25, Application US/08137117D  
; Patent No. 5795965  
; GENERAL INFORMATION:  
; APPLICANT: TSUCHIYA, Masayuki  
; APPLICANT: SATO, Koh  
; APPLICANT: BENDIG, Mary  
; APPLICANT: JONES, Steven  
; APPLICANT: SALDANHA, Jose  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
; NUMBER OF SEQUENCES: 158  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117D  
FILING DATE: 20-DEC-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP92/00544  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-32084  
FILING DATE: 19-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-95476  
FILING DATE: 25-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/126/RAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 131 amino acids  
TYPE: amino acid



US-08-471-426-2  
; Sequence 2, Application US/08471426  
; Patent No. 5808033  
; GENERAL INFORMATION:  
; APPLICANT: GOURLIE, BRIAN B  
; APPLICANT: RIXON, MARK W  
; APPLICANT: MEZES, PETER S  
; APPLICANT: KAPLAN, DONALD A  
; APPLICANT: SCHLOM, JEFFREY  
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC  
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Duane C. Ulmer  
; STREET: P.O. Box 1967  
; CITY: Midland  
; STATE: MI  
; COUNTRY: US  
; ZIP: 48641-1967  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,426  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/017,570  
; FILING DATE: 16-FEB-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ULMER, DUANE C  
; REGISTRATION NUMBER: 34,941  
; REFERENCE/DOCKET NUMBER: C-38,777  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (517) 636-8104  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 110 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-471-426-2

Query Match 88.8%; Score 513.5; DB 1; Length 110;  
Best Local Similarity 90.1%; Pred. No. 3.1e-47;  
Matches 100; Conservative 5; Mismatches 5; Indels 1; Gaps 1;  
  
QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
Db 1 DIVLTQSPASLTVSLGLRATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
  
QY 61 GVPARESGSGGDTFTLNHPVEEDASTYCYQHSGELPFTFGSGTKLEIK 111  
Db 61 GVPARESGSGGDTFTLNHPVEEDASTYCYQHSGELP-TFGGSGTKLEIK 110

RESULT 5  
PCT-US94-01709-2  
; Sequence 2, Application PC/TUS9401709  
; GENERAL INFORMATION:  
; APPLICANT: THE DOW CHEMICAL COMPANY  
; APPLICANT: U.S.A. DEPT. OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC  
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Duane C. Ulmer  
; STREET: P.O. Box 1967  
; CITY: Midland  
; STATE: MI

COUNTRY: US  
ZIP: 48641-1967  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/01709  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: ULMER, DUANE C  
REGISTRATION NUMBER: 34,941  
REFERENCE/DOCKET NUMBER: 38,777-F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 636-8104  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01709-2  
  
Query Match 88.8%; Score 513.5; DB 5; Length 110;  
Best Local Similarity 90.1%; Pred. No. 3.1e-47;  
Matches 100; Conservative 5; Mismatches 5; Indels 1; Gaps 1;  
  
QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
Db 1 DIVLTQSPASLTVSLGLRATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
  
QY 61 GVPARESGSGGDTFTLNHPVEEDASTYCYQHSGELPFTFGSGTKLEIK 111  
Db 61 GVPARESGSGGDTFTLNHPVEEDASTYCYQHSGELP-TFGGSGTKLEIK 110  
  
RESULT 6  
US-07-634-278-46  
; Sequence 46, Application US/07634278  
; Patent No. 5530101  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO, Man Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLFI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/634,278  
; FILING DATE: 19-DEC-1990  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252





TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-474-040-46

Query Match 88.28; Score 510; DB 1; Length 111;  
Best Local Similarity 88.38; Pred. No. 7.2e-47;  
Matches 98; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNYMHVYQKAGQPPKLLIHLASNLES 60  
|||||  
Db 1 DIVLTQSPASLAVSLGORATISCRASQSVSTSYNYMHVYQKAGQPPKLLIKYASNLES 60  
|||||  
QY 61 GVPARFSGSGGTDTLNIHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111  
|||||  
Db 61 GVPARFSGSGGTDTLNIHPVEEDASTYYCQHSWEIPVTFGGGTKLEIK 111  
|||||

RESULT 9  
US-08-487-200-46  
; Sequence 46, Application US/08487200  
; Patent No. 5693762  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO. Map Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLFI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBLINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,200  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/634,278  
; FILING DATE: 19-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252  
; FILING DATE: 13-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/290,975  
; FILING DATE: 28-DEC-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-002610  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400

TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-487-200-46

Query Match 88.28; Score 510; DB 1; Length 111;  
Best Local Similarity 88.38; Pred. No. 7.2e-47;  
Matches 98; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNYMHVYQKAGQPPKLLIHLASNLES 60  
|||||  
Db 1 DIVLTQSPASLAVSLGORATISCRASQSVSTSYNYMHVYQKAGQPPKLLIKYASNLES 60  
|||||  
QY 61 GVPARFSGSGGTDTLNIHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111  
|||||  
Db 61 GVPARFSGSGGTDTLNIHPVEEDASTYYCQHSWEIPVTFGGGTKLEIK 111  
|||||

RESULT 10  
US-08-484-537-46  
; Sequence 46, Application US/08484537  
; Patent No. 6180370  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO. Map Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLFI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBLINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,537  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/634,278  
; FILING DATE: 19-DEC-1990  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252  
; FILING DATE: 13-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/290,975  
; FILING DATE: 28-DEC-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-002600  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-484-537-46

Query Match 88.2%; Score 510; DB 4; Length 111;  
Best Local Similarity 88.3%; Pred. No. 7.2e-47; Indels 0; Gaps 0;  
Matches 98; Conservative 3; Mismatches 10;

Qy 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNYMHVYQKAGQPPKLLIHLASNL 60

Db 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNYMHVYQKAGQPPKLLIHLASNL 60

Qy 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111

Db 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111

## RESULT 11

US-08-207-169A-4

Sequence 4, Application US/08207169A

Patent No. 5674712

GENERAL INFORMATION:

APPLICANT: GRANDI, GUIDO

APPLICANT: DE PERRA, FRANCESCA

APPLICANT: TOSI, CLAUDIO

APPLICANT: TORTORA, ORNELLA

APPLICANT: CUZZONI, ANNA

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

TITLE OF INVENTION: EXOCYTOSED RECOMBINANT VECTOR AND USE THEREOF FOR

Search completed: June 28, 2001, 16:01:15  
Job time: 523 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:58:45 ; Search time 234.85 Seconds  
(without alignments)  
36.003 Million cell updates/sec

Title: US-09-724-406-26  
Perfect score: 578  
Sequence: 1 DIVLTQSPASLAVSLGQRAT.....CQHSGLPFTFGSGTKLEIK 111

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	536	92.7	111	1 KWS75	Ig kappa chain V r
2	534	92.4	111	1 KWS84	Ig kappa chain V r
3	528	91.3	111	1 KWS40	Ig kappa chain V r
4	524	90.7	210	2 A56169	Ig kappa chain V r
5	524	90.7	218	2 JC5810	monoclonal antibody
6	520	90.0	112	2 S45715	Ig kappa chain V r
7	520	90.0	218	2 S68241	Ig kappa chain V r
8	517	89.4	111	1 KWS85	Ig kappa chain V r
9	509	88.1	131	2 S5027	Ig kappa chain V r
10	507	87.7	111	2 S09963	Ig light chain pre
11	493	85.3	111	1 KWS43	Ig kappa chain V-J
12	490	84.8	108	1 KWS54	Ig kappa chain V r
13	488	84.4	111	1 KWS08	Ig kappa chain V r
14	486.5	84.2	110	2 S24288	Ig kappa chain V r
15	486	84.1	140	2 PN0446	Ig kappa chain pre
16	482	83.4	111	1 KWS83	Ig kappa chain pre
17	482	83.4	111	1 KWS69	Ig kappa chain V r
18	479	82.9	112	2 S19972	Ig kappa chain V r
19	476	82.4	102	2 PH1077	Ig light chain V r
20	475	82.2	111	1 KWS01	Ig kappa chain V r
21	473.5	81.9	110	1 KWS10	Ig kappa chain V r
22	473	81.8	112	2 S19976	Ig kappa chain V r
23	471.5	81.6	108	2 PH0092	Ig kappa chain V r
24	469.5	81.2	233	2 JC5322	p53 specific singl
25	468	81.0	112	2 S19971	Ig kappa chain V r
26	468	81.0	131	2 PH1226	Ig kappa chain pre
27	466	80.6	102	2 PH1076	Ig light chain V r
28	461	79.8	101	2 S59640	Ig light chain V r
29	459	79.4	111	1 KWS37	Ig kappa chain V r

30	459	79.4	115	2 S63596	Ig kappa chain V r
31	458	79.2	131	1 KWSM6	Ig kappa chain pre
32	448	77.5	111	2 PL0081	Ig kappa chain V r
33	446.5	77.2	102	2 PC6027	acetylcholine recep
34	446	77.2	111	1 KWS80	Ig kappa chain V r
35	446	77.2	111	1 KWS50	Ig kappa chain V r
36	446	77.2	128	2 S52448	Ig kappa chain V r
37	446	77.2	128	2 JL0073	Ig kappa chain V r
38	444	76.8	111	2 D45722	aberrant kappa tra
39	441	76.3	111	2 S09969	anti-glycoprotein
40	441	76.3	111	2 S53285	Ig kappa chain V-J
41	440	76.1	111	2 S09966	Ig kappa chain V-J
42	439	76.0	107	2 S26343	Ig kappa chain V r
43	437	75.6	132	1 KWS32	Ig kappa chain pre
44	436	75.4	111	2 S09965	Ig kappa chain V-J
45	434	75.1	109	2 PH0093	Ig kappa chain V r

## ALIGNMENTS

RESULT 1

KWS75

Ig kappa chain V region (PC7175) - mouse (tentative sequence)

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 31-Mar-2000

C:Accession: B01938; A01938

R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.

Nature 276, 785-790, 1978

A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.

A:Reference number: A93204; MUID:79073152

A:Accession: B01938

A:Molecule type: protein

A:Residues: 1-111 <WEI>

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:16-94/Domain: immunoglobulin homology <IMM>

F:23-92/Disulfide bonds: #status predicted

Query Match 92.7%; Score 536; DB 1; Length 111;  
Best Local Similarity 91.9%; Pred. No. 1.6e-41;  
Matches 102; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGVNHYHWOQKAGQPPKLLHLASNLES 60

Db 1 DIVLTQSPASLAVSLGQRATISCRASKSVTSVSGSYHMYHWOQKAGQPPKLLIYLASNLES 60

Qy 61 GVPARFSGSGSTDTFTLNHPVEEDASTYYCQHSGLPFTFGSGTKLEIK 111

Db 61 GVPARFSGSGSTDTFTLNHPVEEDASTYYCQHSGLPFTFGAGTKLEK 111

RESULT 2

KWS84

Ig kappa chain V region (PC6684) - mouse (tentative sequence)

C:Species: Mus musculus (house mouse)

C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 31-Mar-2000

C:Accession: A01938

R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.

Nature 276, 785-790, 1978

A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.

A:Reference number: A93204; MUID:79073152

A:Accession: A01938

A:Molecule type: protein

A:Residues: 1-111 <WEI>

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer

F:16-94/Domain: immunoglobulin homology <IMM>

F;23-92/Disulfide bonds: #status predicted

```
Query Match          92.4%; Score 534; DB 1; Length 111;
Best Local Similarity 92.8%; Pred. No. 2.4e-41;
Matches 103; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||
Db 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||

QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||
Db 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||

RESULT 3
KVMS40
Ig kappa chain V region (PC7940) - mouse (tentative sequence)
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 31-Mar-2000
C:Accession: C01938; A01938
R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.
Nature 276, 785-790, 1978
A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.
A:Reference number: A93204; MUID:79073152
A:Accession: C01938
A:Molecule type: protein
A:Residues: 1-111 <WEI>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger superfamily: immunoglobulin V region; immunoglobulin homology
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-94/Domain: immunoglobulin homology <IMM>
F:23-92/Disulfide bonds: #status predicted

Query Match          91.3%; Score 528; DB 1; Length 111;
Best Local Similarity 92.8%; Pred. No. 8.4e-41;
Matches 103; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||
Db 1 DIVLTQSPASLAVSLGORATISCRASKSVSAFGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||

QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||
Db 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||

RESULT 4
A56169
Ig kappa chain V region (clone 23.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 11-Jan-2000
C:Accession: A56169
R:Monfardini, C.; Kleber-Emmons, T.; VonFeltdt, J.M.; O'Malley, B.; Rosenbaum, H.; Godill
J. Biol. Chem. 270, 6628-6638, 1995
A:Title: Recombinant antibodies in bioactive peptide design.
A:Reference number: A56169; MUID:95204454
A:Accession: A56169
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-210 <MON>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match          90.7%; Score 524; DB 2; Length 210;
Best Local Similarity 89.2%; Pred. No. 3.8e-40;
Matches 99; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||
```

```
Db 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||
QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||
Db 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||

RESULT 5
JC5810
monoclonal antibody 13-1 light chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jan-2000
C:Accession: JC5810
R:Akashi, S.; Kato, K.; Torizawa, T.; Dohmae, N.; Yamaguchi, H.; Kamachi, M.; Harada,
Biochem. Biophys. Res. Commun. 240, 566-572, 1997
A:Title: Structural characterization of mouse monoclonal antibody 13-1 against a porp
A:Reference number: JC5810; MUID:98063277
A:Accession: JC5810
A:Molecule type: protein
A:Residues: 1-218 <AKA>
C:Comment: This catalytic antibody has peroxidase oxidase. It is directed against a p
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-94/Domain: immunoglobulin homology <IMM>

Query Match          90.7%; Score 524; DB 2; Length 218;
Best Local Similarity 91.0%; Pred. No. 3.9e-40;
Matches 101; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||
Db 1 NIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||

QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||
Db 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGAGTKLEIK 111
    |||||

RESULT 6
S45715
Ig kappa chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 14-Nov-1997 #text_change 07-May-1999
C:Accession: S45715
R:Kim, H.H.; Kato, K.; Yamato, S.; Igarashi, T.; Matsunaga, C.; Ohtsuka, H.; Higuchi,
FEBS Lett. 346, 246-250, 1994
A:Title: Application of 13C NMR spectroscopy to paratope mapping for larger antigen
A:Reference number: S45714; MUID:94283606
A:Accession: S45715
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-112 <KIM>
A:Experimental source: cell line EX-3C7
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-94/Domain: immunoglobulin homology <IMM>
F:23-92/Disulfide bonds: #status predicted

Query Match          90.0%; Score 520; DB 2; Length 112;
Best Local Similarity 88.3%; Pred. No. 4.5e-40;
Matches 98; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||
Db 1 DIVLTQSPASLAVSLGORATISCRASKSVSGSYNMYHMYQKAGQPPKLLIHLASNL 60
    |||||

QY 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111
    |||||
Db 61 GVPARFSGSGGDTFTLNHPVEEDASTYYCOHSGELPFTFGAGTKLEIK 111
    |||||
```

## RESULT 7

S68241  
 Ig kappa chain V region (Mab13-1) - mouse (fragment)  
 N:Alternate names: Immunoglobulin light chain  
 C:Species: Mus musculus (house mouse)  
 C:Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Jun-2000  
 C:Accession: S68241; S68214  
 R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.  
 submitted to the EMBL Data Library, March 1994  
 A:Description: Specific peroxidase activity by formation of an antibody L-chain-porphyrin  
 A:Reference number: S68241  
 A:Accession: S68241  
 A:Molecule type: mRNA  
 A:Residues: 1-218 <TAK>  
 A:Cross-references: EMBL:D29670; NID:g473962; PIDN:BAA06141.1; PID:g473963  
 R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.  
 FEBS Lett. 375, 273-276, 1995  
 A:Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyrin  
 A:Reference number: S68211; M0ID:96085223  
 A:Accession: S68214  
 A>Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 'NI',3-212 <TAW>  
 A:Cross-references: EMBL:D29670  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: immunoglobulin

	Query Match	90.0%	Score 520;	DB 2;	Length 218;
	Best Local Similarity	89.2%;	Pred. No. 9e-40;		
	Matches	99;	Conservative	6;	Mismatches
				6;	Indels
					0;
					Gaps
QY	1	DIVLTQSPASLAVSLGQRATISCRASKSVSASGYINMHWYQOKAGPPKLLIHLSNLES	60		
	::				
Db	1	EVLVTQSPASLAVSLGQRATISCRASKSVSASGYIYMHWTQKPGPPKLLISLATNLES	60		
QY	61	GVPAFSGSGSGCTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK	111		
Db	61	GVPAFSGSGSGCTDFTLNHPVEEDVATYYCOHSELPLTFGAGTKLEIK	111		

## RESULT 8

KWS85  
 Ig kappa chain V regions (PC2485, PC4039) - mouse (tentative sequence)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 31-Mar-2000  
 C:Accession: A01939  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: A01939  
 A:Molecule type: Protein  
 A:Residues: 1-111 <WEI>  
 A:Note: The PC4285 and PC4039 sequences are identical  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into large C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 89.4%; Score 517; DB 1; Length 111;  
Best Local Similarity 89.2%; pred. NO. 8.2e-40;  
Matches 99; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 61 GVPARFSGSGGTDFTLNIHPVEEDASTYCYQHSGELPTFGSGTKLEIK 111  
|||||.|||||: ||||| ||| |||: |||: |

db 61 GVPARFSGSGTDFTLNIQPVVEEDAAIYYCQHSRELPLTFGAGTKLELK 111

## RESULT 9

S55027  
Ig light chain precursor V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 21-Jan-2000  
C:Accession: S55027  
R:Jeffrey, P.D.; Schildbach, J.F.; Chang, C.Y.; Kussie, P.H.; Margolies, M.N.; Sherif  
J. Mol. Biol. 248, 344-360, 1995  
A:Title: Structure and specificity of the anti-digoxin antibody 40-50.  
A:Reference number: S55027; MUID:95257394  
A:Accession: S55027  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-131 <JEFF>  
A:CROSS-references: EMBL:L31404; NID:g476719; PIDN:AAA72437.1; PID:g476720  
C:Superfamily: Immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:36-114/Domain: immunoglobulin homology <IMM>

Query Match	88.1%;	Score 509;	DB 2;	Length 131;
Best Local Similarity	87.4%;	Pred. No. 5.1e-39;		
Matches 97;	Conservative	8;	Mismatches 6;	Indels 0;
Gaps 0;				

[illegible]

RESULT 10

Ig kappa chain V-J region (103-7E) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 21-Jan-2000  
 C:Accession: S09963  
 R:Reininger, L.; Shibata, T.; Ozaki, S.; Shirai, T.; Jaton, J.C.; Izui, S.  
 Eur. J. Immunol. 20, 771-777, 1990  
 A:Title: Variable region sequences of pathogenic anti-mouse red blood cell autoantibody  
 A:Reference number: S09955; MUID:90269328  
 A:Accession: S09963  
 A:Molecule type: mRNA  
 A:Residues: 1-111 <REI>  
 A:Cross-references: EMBL:X51851  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:16-94/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 507; DB 2; Length 111;  
Best Local Similarity 86.5%; Pred. No. 6.5e-39;  
Matches 96; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

**QY** 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
|||  
|||  
**Dd** 61 GVPARFSGSGGTDFTLNHPVEEDATYYCOHSWEIPYTFGGTKLEIK 111

RESULT 11

II  
RESULTS  
KVMS43  
Ig kappa chain V region (PC7043) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:date: 01-Sep-1981 #sequence\_revision 01-Sep-1981

C:Accession: A01937; S42187; S42194; S42190; S42189; S42188; S42191; S42192  
 R:Reigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: A01937  
 A:Molecule type: protein  
 A:Residues: 1-111 <WEI>  
 R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
 Eur. J. Immunol. 23, 2503-2510, 1993  
 A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with spec  
 A:Reference number: S42187  
 A:Accession: S42187  
 A:Molecule type: DNA  
 A:Residues: 10-99 <MOJ>  
 A:Cross-references: EMBL:Z25444; NID:g407832; PIDN:CAA80931.1; PID:g407833  
 A:Note: V-kappa-2IE; anti-collagen  
 A:Accession: S42194  
 A:Molecule type: DNA  
 A:Residues: 12-99 <NOW>  
 A:Cross-references: EMBL:Z25458; NID:g407844; PIDN:CAA80945.1; PID:g407845  
 A:Note: V-kappa-2IE; anti-collagen  
 A:Accession: S42190  
 A:Molecule type: DNA  
 A:Residues: 13-99 <MOF>  
 A:Cross-references: EMBL:Z25450; NID:g407838; PIDN:CAA80937.1; PID:g407839  
 A:Note: V-kappa-2IE; anti-collagen  
 A:Accession: S42189  
 A:Molecule type: DNA  
 A:Residues: 15-99 <MOA>  
 A:Cross-references: EMBL:Z25448; NID:g407836; PIDN:CAA80935.1; PID:g407837  
 A:Note: V-kappa-2IE; anti-collagen  
 A:Accession: S42188  
 A:Molecule type: DNA  
 A:Residues: 12-99 <MOZ>  
 A:Cross-references: EMBL:Z25446; NID:g407834; PIDN:CAA80933.1; PID:g407835  
 A:Note: V-kappa-2IE; anti-collagen  
 A:Accession: S42191  
 A:Molecule type: DNA  
 A:Residues: 10-99 <MOY>  
 A:Cross-references: EMBL:Z25452; NID:g407840; PIDN:CAA80939.1; PID:g407841  
 A:Note: V-kappa-2IE; anti-collagen  
 A:Accession: S42192  
 A:Molecule type: DNA  
 A:Residues: 10-99 <MOO>  
 A:Cross-references: EMBL:Z25454; NID:g407842; PIDN:CAA80941.1; PID:g407843  
 A:Note: V-kappa-2IE; anti-collagen  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 85.3%; Score 493; DB 1; Length 111;  
 Best Local Similarity 85.6%; Pred. No. 1.2e-37;  
 Matches 95; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGQPPKLLIHLASNLES 60  
 |||||  
 Db 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGQPPKLLIHLASNLES 60  
 |||||

Qy 61 GVPARFSGSGTDFTLNIHPVEEDASTYYCQHSQSGELPFTFGSGTKLEIK 111  
 |||||  
 Db 61 GIPARFSGSGTDFTLNIHPVEEDASTYYCQHSQSGELPFTFGSGTKLEIK 111  
 |||||

RESULT 12  
 KVM54  
 Ig kappa chain v region (PC2154) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 21-Jan-2000

C:Accession: A01940  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: A01940  
 A:Molecule type: protein  
 A:Residues: 1-108 <WEI>  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (h  
 ain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 84.8%; Score 490; DB 1; Length 108;  
 Best Local Similarity 86.1%; Pred. No. 2.1e-37;  
 Matches 93; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGQPPKLLIHLASNLES 60  
 |||||  
 Db 1 DIVLAQSPASLTVSLGQRATISCRASQSVSTSGYSNMHWYQKPGQPPKLLIKYASNLES 60  
 |||||

Qy 61 GVPARFSGSGTDFTLNIHPVEEDASTYYCQHSQSGELPFTFGSGTKL 108  
 |||||  
 Db 61 GVPARFSGSGTDFTLNIHPVEEDASTYYCQHSQSGELPFTFGAGTKL 108  
 |||||

RESULT 13  
 KVM508  
 Ig kappa chain V region (PC6308) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 21-Jan-2000  
 C:Accession: C01937; A01937  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: C01937  
 A:Molecule type: protein  
 A:Residues: 1-111 <WEI>  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (h  
 ain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 84.4%; Score 488; DB 1; Length 111;  
 Best Local Similarity 84.7%; Pred. No. 3.3e-37;  
 Matches 94; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGQPPKLLIHLASNLES 60  
 |||||  
 Db 1 DIVLTQSPASLAVSLGQRATISCRASQSVSDYDGSYMNWYQKPGQPPKLLIYASNLES 60  
 |||||

Qy 61 GVPARFSGSGTDFTLNIHPVEEDASTYYCQHSQSGELPFTFGSGTKLEIK 111  
 |||||  
 Db 61 GIPARFSGSGTDFTLNIHPVEEDASTYYCQHSQSGELPFTFGSGTKLEIK 111  
 |||||

RESULT 14  
 S24288  
 Ig kappa chain V region (JS28/32) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000  
 C:Accession: S24288  
 R:Nonchamont, B.  
 submitted to the EMBL Data Library, September 1991  
 A:Description: Cloning and sequencing of the cDNA coding for the variable regions of  
 A:Reference number: S24287



A:Accession: S24288  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-110 <MON>  
A:Cross-references: EMBL:X62703; NID:G51673; PIDN:CAA44576.1; PID:gl333958  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:16-94/Domain: immunoglobulin homology <IMM>

Query Match 84.2%; Score 486.5; DB 2; Length 110;  
Best Local Similarity 86.5%; Pred. No. 4.5e-37;  
Matches 96; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVSYSGYNYMHYQKAGQPPKLLIHLASNL 60

DB 1 DIQLTQSPASLAVSLGQRATISYRASKSVSYSGYNYMHYQKAGQPPKLLIHLASNL 60

QY 61 GVPARFSGSGCTDFTLNHPVEEDASTYCYQHSGELPFTFGSGTKLEIK 111

DB 61 GVPARFSGSGCTDFTLNHPVEEDASTYCYQHSGELPFTFGSGTKLEIK 110

## RESULT 15

PN0446

Ig kappa chain precursor V-II region - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000

C:Accession: PN0446

R:Kaluza, B.; Betzl, G.; Shao, H.; Diamantstein, T.; Weidle, U.H.

Gene 122, 321-328, 1992

A:Title: A general method for chimerization of monoclonal antibodies by inverse polymera

A:Reference number: PN0444; MUID:93138402

A:Accession: PN0446

A:Molecule type: mRNA

A:Residues: 1-140 &lt;KAL&gt;

A:Cross-references: GB:L02345

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-20/Domain: signal sequence #status predicted &lt;SIG&gt;

F:21-140/Product: Ig light chain kappa-2 V region #status predicted &lt;MAT&gt;

F:36-114/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 84.1%; Score 486; DB 2; Length 140;  
Best Local Similarity 87.4%; Pred. No. 6.4e-37;  
Matches 97; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVSYSGYNYMHYQKAGQPPKLLIHLASNL 60

DB 21 DIVLTQSPASLAVSLGQRATISYRASKSVSYSGYNYMHYQKAGQPPKLLIHLASNL 80

QY 61 GVPARFSGSGCTDFTLNHPVEEDASTYCYQHSGELPFTFGSGTKLEIK 111

DB 81 GVPARFSGSGCTDFTLNHPVEEDASTYCYQHREL--TFGGGKLEIK 129

Search completed: June 28, 2001, 15:58:46  
Job time: 374 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:54:38 ; Search time 105.36 Seconds  
(without alignments)  
36.089 Million cell updates/sec

Title: US-09-724-406-26

Perfect score: 578

Sequence: 1 DIVLTQSPASLAVSLGQRAT.....COHSGELPFTFGSGTKLEIK 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	536	92.7	111	1	KV3S_MOUSE
2	534	92.4	111	1	KV3R_MOUSE
3	528	91.3	111	1	KV3T_MOUSE
4	517	89.4	111	1	KV3U_MOUSE
5	493	85.3	111	1	KV3M_MOUSE
6	490	84.8	108	1	KV3V_MOUSE
7	488	84.4	111	1	KV3O_MOUSE
8	482	83.4	111	1	KV3N_MOUSE
9	482	83.4	111	1	KV3Q_MOUSE
10	475	82.2	111	1	KV3L_MOUSE
11	473.5	81.9	110	1	KV3P_MOUSE
12	460	79.6	111	1	KV3J_MOUSE
13	459	79.4	111	1	KV3H_MOUSE
14	458	78.2	131	1	KV3I_MOUSE
15	456	78.9	111	1	KV3K_MOUSE
16	452	78.2	111	1	KV3C_MOUSE
17	446	77.2	111	1	KV3D_MOUSE
18	446	77.2	111	1	KV3A_MOUSE
19	437.5	75.7	112	1	KV3B_MOUSE
20	437	75.6	132	1	KV3F_MOUSE
21	433	74.9	112	1	KV3G_MOUSE
22	418	72.3	111	1	KV3E_MOUSE
23	385	66.6	134	1	KV4C_HUMAN
24	383	66.3	114	1	KV4A_HUMAN
25	374.5	64.8	133	1	KV4B_HUMAN
26	369	63.8	108	1	KV5P_MOUSE
27	359	62.1	129	1	KV1W_HUMAN
28	352.5	61.0	129	1	KV3L_HUMAN
29	351.5	60.8	129	1	KV3M_HUMAN
30	350.5	60.6	109	1	KV3D_HUMAN
31	349.5	60.5	113	1	KV2G_MOUSE
32	349.5	60.5	129	1	KV4A_MOUSE
33	347	60.0	108	1	KV1B_HUMAN

34	347	60.0	108	1	KV1Y_HUMAN	P80362 homo sapien
35	346.5	59.9	109	1	KV3E_HUMAN	P01623 homo sapien
36	346.5	59.9	113	1	KV2E_MOUSE	P03976 mus musculus
37	345.5	59.8	112	1	KV2D_MOUSE	P01629 mus musculus
38	345	59.7	108	1	KV1V_HUMAN	P04430 homo sapien
39	344.5	59.6	117	1	KV2E_HUMAN	P06309 homo sapien
40	343.5	59.4	109	1	KV3B_HUMAN	P01620 homo sapien
41	342.5	59.3	113	1	KV2D_HUMAN	P01617 homo sapien
42	342.5	59.3	129	1	KV3H_HUMAN	P04207 homo sapien
43	342	59.2	108	1	KV1H_HUMAN	P01600 homo sapien
44	342	59.2	108	1	KV1K_HUMAN	P01603 homo sapien
45	342	59.2	108	1	KV1M_HUMAN	P01605 homo sapien

## ALIGNMENTS

RESULT 1						
KV3S_MOUSE						
ID KV3S_MOUSE	STANDARD;	PRT;	111 AA.			
AC P01671;						
DT 21-JUL-1986 (Rel. 01, Created)						
DT 21-JUL-1986 (Rel. 01, Last sequence update)						
DT 15-JUL-1999 (Rel. 38, Last annotation update)						
DE IG KAPPA CHAIN V-III REGION PC 7175.						
OS Mus musculus (Mouse).						
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
OX NCBI_TaxID=10090;						
RN [1]						
RP SEQUENCE.						
RX MEDLINE=79073152; PubMed=103003;						
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;						
RT "Rearrangement of genetic information may produce immunoglobulin						
RT diversity."						
RL Nature 276:785-790(1978).						
DR PIR; B01938; KVM575.						
DR InterPro; IPR003006;						
DR Pfam; PF00047; Ig; 1.						
KW Immunoglobulin V region.						
FT DOMAIN 1 23						
FT DOMAIN 24 38						
FT DOMAIN 39 53						
FT DOMAIN 54 60						
FT DOMAIN 61 92						
FT DOMAIN 93 101						
FT DOMAIN 102 111						
FT DISULFID 23 92						
FT NON_TER 111 111						
SQ SEQUENCE 111 AA; 12010 MW; F041E89AA7858523 CRC64;						
Query Match	92.7%;	Score 536;	DB 1;	Length 111;		
Best Local Similarity	91.9%;	Pred. No. 8.5e-48;				
Matches 102;	Conservative 5;	Mismatches 4;	Indels 0;	Gaps 0;		
Qy 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNHWHYQKAGQPPKLLIHLASNL	60					
Db 1 DIVLTQSPASLAVSLGQRATISCRASKSVTSYSTSYHWHYQKQPPKLLIHLASNL	60					
Qy 61 GVPARFSGSGSDFTLNHPVEEEDASTYVCOHSGELPFTFGSGTKLEIK 111						
Db 61 GVPARFSGSGSDFTLNHPVEEEDASTYVCOHSGELPFTFGAGTKLEIK 111						
RESULT 2						
KV3R_MOUSE						
ID KV3R_MOUSE	STANDARD;	PRT;	111 AA.			
AC P01670;						
DT 21-JUL-1986 (Rel. 01, Created)						
DT 21-JUL-1986 (Rel. 01, Last sequence update)						
DT 15-JUL-1999 (Rel. 38, Last annotation update)						
DE IG KAPPA CHAIN V-III REGION PC 6684.						

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 diversity";  
 RL Nature 276:785-790(1978).  
 DR PIR: A01938; KVM584.  
 DR InterPro: IPR003006;  
 DR Pfam: PF00047; Ig; 1.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12039 MW; 1E46988341858526 CRC64;

Query Match 92.4%; Score 534; DB 1; Length 111;  
 Best Local Similarity 92.8%; Pred. No. 1.4e-47;  
 Matches 103; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 DB 1 DIVLTQSPASLAVSLGQRATISCRASKSVTSYGYNMHWYQKAGOPPKLLIHLASNL 60  
 QY 61 GVPARFSGSGGTDTLNIHPVEEDASTYCYQHSGLPFTFGSGTKLEIK 111  
 DB 61 GVPARFSGSGGTDTLNIHPVEEDASTYCYQHSRELPTFGGKLEIK 111

RESULT 3  
 KV3T\_MOUSE  
 ID KV3T\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01672;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 7940.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 diversity";  
 RL Nature 276:785-790(1978).  
 DR PIR: C01938; KVM540.  
 DR InterPro: IPR003006;  
 DR Pfam: PF00047; Ig; 1.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12038 MW; E88A82306084352E CRC64;

Query Match 91.3%; Score 528; DB 1; Length 111;  
 Best Local Similarity 92.8%; Pred. No. 5.5e-47;  
 Matches 103; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 DB 1 DIVLTQSPASLAVSLGQRATISCRASKSVSAFYGYNMHWYQKAGOPPKLLIHLASNL 60  
 QY 61 GVPARFSGSGGTDTLNIHPVEEDASTYCYQHSGLPFTFGSGTKLEIK 111  
 DB 61 GVPARFSGSGGTDTLNIHPVEEDASTYCYQHSRELPTFGGKLEIK 111

RESULT 4  
 KV3U\_MOUSE  
 ID KV3U\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01673;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 2485/PC 4039.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 diversity";  
 RL Nature 276:785-790(1978).  
 CC -1- MISCELLANEOUS: THE PC 4285 AND PC 4039 SEQUENCES ARE IDENTICAL.  
 DR PIR: A01939; KVM585.  
 DR InterPro: IPR003006;  
 DR Pfam: PF00047; Ig; 1.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 11986 MW; BF38C59AA7858467 CRC64;

Query Match 89.4%; Score 517; DB 1; Length 111;  
 Best Local Similarity 89.2%; Pred. No. 7.2e-46;  
 Matches 99; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 DB 1 DIVLTQSPASLAVSLGQRATISCRASKSVTSYGYNMHWYQKAGOPPKLLIHLASNL 60  
 QY 61 GVPARFSGSGGTDTLNIHPVEEDASTYCYQHSGLPFTFGSGTKLEIK 111  
 DB 61 GVPARFSGSGGTDTLNIHPVEEDASTYCYQHSRELPTFGGKLEIK 111

RESULT 5  
 KV3M\_MOUSE  
 ID KV3M\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01685;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 7043.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 RT diversity.";  
 RL Nature 276:785-790(1978).  
 DR PIR: A01937; KVM543.  
 DR InterPro: IPR003006; -.  
 DR Pfam: PF00047; Ig: 1.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111  
 SQ SEQUENCE 111 AA; 12002 MW; 7A5FCB586C306D29 CRC64;

Query Match 85.3%; Score 493; DB 1; Length 111;  
 Best Local Similarity 85.6%; Pred. No. 1.9e-43;  
 Matches 95; Conservative 7; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 Db 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 QY 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
 Db 61 GIPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111

RESULT 6  
 KV3V\_MOUSE  
 ID KV3V\_MOUSE STANDARD; PRT; 108 AA.  
 AC P01674;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 2154.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 RT diversity.";  
 RL Nature 276:785-790(1978).  
 DR PIR: A01940; KVM54.  
 DR InterPro: IPR003006; -.  
 DR Pfam: PF00047; Ig: 1.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 108  
 FT DISULFID 23 92  
 FT NON\_TER 108  
 SQ SEQUENCE 108 AA; 11699 MW; D40921D18DAC4B9E CRC64;

Query Match 84.8%; Score 490; DB 1; Length 108;  
 Best Local Similarity 86.1%; Pred. No. 3.8e-43;  
 Matches 93; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 Db 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 QY 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKL 108  
 Db 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGAGTKL 108

RESULT 7  
 KV3O\_MOUSE  
 ID KV3O\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01667;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 6308.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 RT diversity.";  
 RL Nature 276:785-790(1978).  
 DR PIR: C01937; KVM508.  
 DR InterPro: IPR003006; -.  
 DR Pfam: PF00047; Ig: 1.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111  
 SQ SEQUENCE 111 AA; 12071 MW; 7A4ADE4D6C256D29 CRC64;

Query Match 84.4%; Score 488; DB 1; Length 111;  
 Best Local Similarity 84.7%; Pred. No. 6.3e-43;  
 Matches 94; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 Db 1 DIVLTQSPASLAVSLGORATISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNL 60  
 QY 61 GVPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
 Db 61 GIPARFSGSGGTDFTLNHPVEEDASTYYCOHSGELPFTFGSGTKLEIK 111

RESULT 8  
 KV3N\_MOUSE  
 ID KV3N\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01666;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 7183.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
diversity.";
RL Nature 276:785-790(1978).
DR PIR; B01937; KVM583.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111
SQ SEQUENCE 111 AA; 11952 MW; 2058BBS50CE306D31 CRC64;

Query Match 83.4%; Score 482; DB 1; Length 111;
Best Local Similarity 82.9%; Pred. No. 2.5e-42;
Matches 92; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Qy 1 DIVLTQSPASVLSLQGRATISCRASKSVASGYNMYHWYQKAGOPPKLLIHLASNLDES 60
Db 1 DIVLTQSPASVLSLQGRATISCRASKSVASGYNMYHWYQKAGOPPKLLIHLASNLDES 60

Qy 61 GVPARFSGSGGTDFTLNHPVEEEDASTYCYQHSGLPFTFGSGTKLEIK 111
Db 61 GIPARFSGSGGTDFTLNHPVEEEDASTYCYQHSGLPFTFGSGTKLEIK 111

RESULT 9
KV3Q_MOUSE STANDARD; PRT; 111 AA.
ID KV3Q_MOUSE
AC P01669;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG KAPPA CHAIN V-III REGION PC 7769.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

Query Match 83.4%; Score 482; DB 1; Length 111;
Best Local Similarity 82.9%; Pred. No. 2.5e-42;
Matches 92; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Qy 1 DIVLTQSPASVLSLQGRATISCRASKSVASGYNMYHWYQKAGOPPKLLIHLASNLDES 60
Db 1 DIVLTQSPASVLSLQGRATISCRASKSVASGYNMYHWYQKAGOPPKLLIHLASNLDES 60

Qy 61 GVPARFSGSGGTDFTLNHPVEEEDASTYCYQHSGLPFTFGSGTKLEIK 111
Db 61 GIPARFSGSGGTDFTLNHPVEEEDASTYCYQHSGLPFTFGSGTKLEIK 111

RESULT 10
KV3L_MOUSE STANDARD; PRT; 111 AA.
ID KV3L_MOUSE
AC P01664;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG KAPPA CHAIN V-III REGION CBPC 101.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

Query Match 82.2%; Score 475; DB 1; Length 111;
Best Local Similarity 82.0%; Pred. No. 1.3e-41;
Matches 91; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Qy 1 DIVLTQSPASVLSLQGRATISCRASKSVASGYNMYHWYQKAGOPPKLLIHLASNLDES 60
Db 1 DIVLTQSPASVLSLQGRATISCRASKSVASGYNMYHWYQKAGOPPKLLIHLASNLDES 60

Qy 61 GVPARFSGSGGTDFTLNHPVEEEDASTYCYQHSGLPFTFGSGTKLEIK 111
Db 61 GIPARFSGSGGTDFTLNHPVEEEDASTYCYQHSGLPFTFGSGTKLEIK 111

RESULT 11
KV3P_MOUSE STANDARD; PRT; 110 AA.
ID KV3P_MOUSE
AC P01668;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG KAPPA CHAIN V-III REGION PC 7210.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```



## RESULT 14

KV31\_MOUSE STANDARD; PRT; 131 AA.  
 AC P01661;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION MOPC 63 PRECURSOR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE OF 1-35.  
 RX MEDLINE=78235887; PubMed=98179;  
 RA Burstein Y., Schechter I.;  
 RT "Primary structures of N-terminal extra peptide segments linked to  
 the variable and constant regions of immunoglobulin light chain  
 precursors: implications on the organization and controlled  
 expression of immunoglobulin genes.";  
 RL Biochemistry 17:2392-2400(1978).  
 RN [2]  
 RP SEQUENCE OF 21-131.  
 RX MEDLINE=73140225; PubMed=4691517;  
 RA McKean D.J., Potter M., Hood L.E.;  
 RT "Mouse immunoglobulin chains. Pattern of sequence variation among  
 kappa chains with limited sequence differences.";  
 RL Biochemistry 12:760-771(1973).  
 RN [3]  
 RP REVISIONS.  
 RX MEDLINE=79012520; PubMed=99744;  
 RA McKean D.J., Bell M., Potter M.;  
 RT "Mechanisms of antibody diversity: multiple genes encode structurally  
 related mouse kappa variable regions.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).  
 DR PIR; A01935; KVM5M6.  
 DR InterPro: IPR003006; -.  
 DR Pfam; PF00047; ig; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 131 IG KAPPA CHAIN V-III REGION MOPC 63.  
 FT DOMAIN 21 43 FRAMEWORK 1.  
 FT DOMAIN 44 58 COMPLEMENTARITY-DETERMINING 1.  
 FT DOMAIN 59 73 FRAMEWORK 2.  
 FT DOMAIN 74 80 COMPLEMENTARITY-DETERMINING 2.  
 FT DOMAIN 81 112 FRAMEWORK 3.  
 FT DOMAIN 113 121 COMPLEMENTARITY-DETERMINING 3.  
 FT DOMAIN 122 131 FRAMEWORK 4.  
 FT DISULFID 43 112 BY SIMILARITY.  
 FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 14291 MW; D212EC9F08DC880A CRC64;

Query Match 79.2%; Score 458; DB 1; Length 131;  
 Best Local Similarity 80.2%; Pred. No. 8.3e-40;  
 Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGQPPKLLIHLASNL 50  
 :|||||  
 Db 21 NVLTQSPASLAVSLGQRATISCRASEVDSYGNFPMHWYQKAGQPPKLLIHLASNL 80  
 :|||||  
 QY 61 GVPARFSGSGGTDFTLTINHPVEEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
 :|||||  
 Db 81 GVPARFSGSGSRDFTLTIDPVEADDAATYYCQNNEDPWTFGGKLEIK 131  
 :|||||

## RESULT 15

KV3K\_MOUSE STANDARD; PRT; 111 AA.  
 ID KV3K\_MOUSE  
 AC P01663;  
 DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 4050.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 diversity.";  
 RL Nature 276:785-790(1978).  
 DR PIR; A01935; KVM5M6.  
 DR InterPro: IPR003006; -.  
 DR Pfam; PF00047; ig; 1.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38 FRAMEWORK 1.  
 FT DOMAIN 39 53 FRAMEWORK 2.  
 FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING 2.  
 FT DOMAIN 61 92 FRAMEWORK 3.  
 FT DOMAIN 93 101 COMPLEMENTARITY-DETERMINING 3.  
 FT DOMAIN 102 111 FRAMEWORK 4.  
 FT DISULFID 23 92 BY SIMILARITY.  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12005 MW; 39D87619313453CB CRC64;

Query Match 78.9%; Score 456; DB 1; Length 111;  
 Best Local Similarity 79.3%; Pred. No. 1.1e-39;  
 Matches 88; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNMHWYQKAGQPPKLLIHLASNL 60  
 :|||||  
 Db 1 NVLTQSPASLAVSLGQRATISCRASEVDSYGNFPMHWYQKAGQPPKLLIHLASNL 60  
 :|||||  
 QY 61 GVPARFSGSGGTDFTLTINHPVEEEDASTYYCOHSGELPFTFGSGTKLEIK 111  
 :|||||  
 Db 61 GVPARFSGSGSRDFTLTIDPVEADDAATYYCQNNEDPWTFGGKLEIK 111  
 :|||||

Search completed: June 28, 2001, 15:54:39  
 Job time: 127 sec



Fri Jun 29 08:04:57 2001

us-09-724-406-26.rsp

---

Page 7



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:08:26 ; Search time 411.58 Seconds  
(without alignments)  
35.682 Million cell updates/sec

Title: US-09-724-406-26  
Perfect score: 578  
Sequence: 1 DIVLTQSPASLAVSIGQRAT.....CQHSGLPFTFGSGTKLEIK 111

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	413	71.5	103	11	Q9JL80	Q9JL80 mus musculu	
2	358.5	62.0	109	4	Q9UL78	Q9ul78 homo sapien	
3	357	61.8	108	4	Q9UL77	Q9ul77 homo sapien	
4	356	61.6	108	4	Q9UL70	Q9ul70 homo sapien	
5	355	61.4	108	4	Q9UL83	Q9ul83 homo sapien	
6	343.5	59.4	107	4	Q9UL81	Q9ul81 homo sapien	
7	343	59.3	108	4	Q9UL79	Q9ul79 homo sapien	
8	337	58.3	298	11	Q9QVF0	Q9qvf0 mus musculu	
9	332	57.4	107	11	Q9ER29	Q9er29 mus musculu	
10	330.5	57.2	109	4	Q9UL86	Q9ul86 homo sapien	
11	328	56.7	214	11	Q9RIAS	Q9ria5 mus musculu	
12	326.5	56.5	104	11	Q9JL82	Q9Jl82 mus musculu	
13	326.5	56.5	106	5	Q9U410	Q9u410 schistosoma	
14	326	56.4	114	4	Q9UL80	Q9ul80 homo sapien	
15	320.5	55.4	109	4	Q9UL85	Q9ul85 homo sapien	
16	316	54.7	101	11	Q9JL78	Q9Jl78 mus musculu	
17	309	53.5	97	11	Q9JL76	Q9Jl76 mus musculu	
18	299	51.7	107	11	Q9JL84	Q9Jl84 mus musculu	
19	296	51.2	109	6	Q9N0W5	Q9n0w5 oryctolagus	

20	292	50.5	99	11	Q9JL74	Q9JL74 mus musculus
21	203	35.1	107	4	Q9UL82	Q9UL82 homo sapien
22	201.5	34.9	107	4	Q9NSD6	Q9NSD6 homo sapien
23	194	33.6	130	4	Q9NP29	Q9NP29 homo sapien
24	145.5	25.2	109	11	Q9ET13	Q9ET13 mus musculus
25	129	22.3	135	4	Q9H524	Q9H524 homo sapien
26	127	22.0	93	4	Q9UL76	Q9UL76 homo sapien
27	116.5	20.2	337	13	Q9IB02	Q9IB02 spherooides
28	116	20.1	123	11	Q61243	Q61243 mus musculus
29	112	19.4	168	4	Q9UQ56	Q9UQ56 homo sapien
30	111.5	19.3	185	4	Q95775	Q95775 homo sapien
31	107.5	18.6	122	4	Q99603	Q99603 homo sapien
32	107.5	18.6	137	4	Q9UDR1	Q9UDR1 homo sapien
33	106.5	18.4	122	4	Q99604	Q99604 homo sapien
34	105.5	18.3	334	13	Q9IB05	Q9IB05 spherooides
35	104	18.0	134	13	Q9YHF8	Q9YHF8 ginglymosto
36	103.5	17.9	141	11	Q9R041	Q9R041 meriones un
37	103	17.8	113	11	Q9ESQ0	Q9ESQ0 rattus norv
38	101	17.5	109	11	Q9JL85	Q9JL85 mus musculus
39	101	17.5	145	4	Q9H022	Q9H022 homo sapien
40	101	17.5	342	13	Q9IB00	Q9IB00 spherooides
41	101	17.5	504	4	Q9UIJ6	Q9UIJ6 homo sapien
42	101	17.5	1021	4	Q93033	Q93033 homo sapien
43	100.5	17.4	235	6	Q9XSM6	Q9XSM6 salmisi sci
44	100	17.3	152	13	Q9YH11	Q9YH11 ginglymosto
45	100	17.3	506	6	Q46631	Q46631 bos taurus

## ALIGNMENTS

RESULT 1  
Q9JL80  
ID Q9JL80 PRELIMINARY: PRT; 103 AA.  
AC Q9JL80;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE ANTI-MYOSIN IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION (FRAGMENT).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C;  
RA Malkiel S., Liao L., Cunningham M.W., Diamond B.;  
RT "Characterization of cross-reactive monoclonal anti-myosin/anti-n-  
acetyl-glucosamine antibodies from mice with autoimmune myocarditis.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF206026; AAF69324.1; -  
DR InterPro; IPR003006; -  
DR InterPro; IPR003596; -  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGv; 1.  
FT NON\_TER 1  
FT NON\_TER 103  
SQ SEQUENCE 103 AA; 11224 MW; EC87D653DB3AAB21 CRC64;

Query Match 71.5%; Score 413; DB 11; Length 103;  
Best Local Similarity 76.7%; Pred. No. 5.5e-38;  
Matches 79; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

Qy 9 ASLAVSLGORATFISCRASKSVASGYNMHWYQKAGOPPKLLIHLASNLGSGVPARFSG 68

Db 1 ASLAVSLGORATFISCRASEVEYGTSLMQWYQKQKOPPKLLIYAASNVESGVPARFSG 60

Qy 69 SGSGTDFTLNLPVEEDASTYCYCOHSGELPFTFGSGTKLEIK 111

Db 61 SGSGTDFSLNLPVEEDDIAMVFCQSRKVPWTFGGTKLEIK 103

```

RESULT 3
ID Q9UL77 PRELIMINARY; PRT; 108 AA.
AC Q9UL77;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RX [1]
RN .SEQUENCE FROM N.A.
RP MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF035037; AAD56273.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; -.
DR InterPro; IPR003596; -.
DR Pfam; PF00047; Iq; 1.

```

	Query Match	61.6%	Score 356;	DB 4;	Length 108;
	Best Local Similarity	62.2%;	pred. No. 1.le-31;		
	Matches 69;	Conservative 15;	Mismatches 23;	Indels 4;	Gaps 1
QY	1	DIVLTQSPASLAVSLGQRATISCRASKSVSASGYNYMHVYQOKAQOPPKLLIHLSANLES	60		
Db	1	DIQMTQSPSSLSASVYGVDRITITCRASQGIS----	NYLAWYQOKPKVPSKLSIYAAASTLQS	56	
QY	61	GVPAFPGSGSGCTDTFLNIHPVEEEDASTYYCQHSGELPFTFGSGTKLEIK	111		
Db	57	GVPSFPGSGSGCTDTFLATSSLOPEDVATYYCOKYNSAPRTFGPGTKLEIK	107		

RESULT	5
Q9UL83	
ID	Q9UL83
AC	PRELIMINARY; PRT; 108 AA.
DT	Q9UL83;
DT	01-MAY-2000 (TREMblrel. 13, Created)
DT	01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT	01-MAY-2001 (TREMblrel. 16, Last annotation update)
DE	MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION (FRAGMENT)

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF035031; AAD56267.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; -.
DR InterPro; IPR003596; -.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11834 MW; 9F9C5A92EBA96EEA CRC64;

Query Match 61.4%; Score 355; DB 4; Length 108;
Best Local Similarity 61.3%; Pred. No. 1.4e-31;
Matches 68; Conservative 20; Mismatches 19; Indels 4; Gaps 1;

QY 1 DIVLTQSPASVSLGORATISCRASKSVASGYNYMHYQKAGOPPKLLIHLASNL 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 EIVNTQSPATLSVPGERATISCRASQSSN---LAWYQKPGQAPRLIYCASTRAT 56
QY 61 GVPARFSGSGGTDFTLNIHPVEEDASTYCOHSGELPFTFGSGTKLEIK 111
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 GIPARFSGSGGTDFTLTIISLQAFDFAVYCYQHYNNWPTFGTKVDIK 107

RESULT 6
QYUL81 PRELIMINARY; PRT; 107 AA.
ID Q9UL81
AC Q9UL81
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF035033; AAD56269.1; -.
DR HSSP; P80362; IWTL.
DR InterPro; IPR003006; -.
DR InterPro; IPR003596; -.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;

Query Match 59.4%; Score 343.5; DB 4; Length 107;
Best Local Similarity 60.4%; Pred. No. 2.5e-30;

```

```

Matches 67; Conservative 20; Mismatches 19; Indels 5; Gaps 2;
QY 1 DIVLTQSPASVSLGORATISCRASKSVASGYNYMHYQKAGOPPKLLIHLASNL 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 DIQWTQSPSSLSASVGDVITICRASQSSIS---NYLWYQKPGKAPNLLIYAASLQS 56
QY 61 GVPARFSGSGGTDFTLNIHPVEEDASTYCOHSGELPFTFGSGTKLEIK 111
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 GVPARFSGSGGTDFTLTIISLQAFDFAVYCOQSSYS-ALTFPGTKVDIR 106

RESULT 7
QYUL79 PRELIMINARY; PRT; 108 AA.
ID Q9UL79
AC Q9UL79
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF035035; AAD56271.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; -.
DR InterPro; IPR003596; -.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11787 MW; DB5845F19724FB4E CRC64;

Query Match 59.3%; Score 343; DB 4; Length 108;
Best Local Similarity 60.4%; Pred. No. 2.8e-30;
Matches 67; Conservative 16; Mismatches 24; Indels 4; Gaps 1;

QY 1 DIVLTQSPASVSLGORATISCRASKSVASGYNYMHYQKAGOPPKLLIHLASNL 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 DIVMTQSPSLLSASTGDRVTISCRMSQGIS---SYLAWYQKPGKAPPELLIYAASLQS 56
QY 61 GVPARFSGSGGTDFTLNIHPVEEDASTYCOHSGELPFTFGSGTKLEIK 111
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 GVPARFSGSGGTDFTLTIISLQAFDFAVYCYQHYNNWPTFGTKVDIR 107

RESULT 8
QYUL79 PRELIMINARY; PRT; 298 AA.
ID Q9YU79
AC Q9YU79
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C; TISSUE=SPLEEN;

```

RA Shinohara N., Demura T., Fukuda H.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C; TISSUE=SPLEEN;  
 RA Shinohara N., Demura T., Fukuda H.;  
 RT Isolation of a novel type of vascular cell wall-specific monoclonal  
 RT antibody recognizing a cell polarity using a phase display subtraction  
 RT method.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB036341; BAA86633.1; -;  
 DR HSP; P01607; IREI.  
 DR InterPro: IPR003006; -;  
 DR InterPro: IPR003596; -;  
 DR Pfam: PF00047; Ig; 2.  
 DR SMART; SM00406; IGV; 1.  
 SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 58.3%; Score 337; DB 11; Length 298;  
 Best Local Similarity 56.8%; Pred. No. 4.3e-29;  
 Matches 63; Conservative 20; Mismatches 24; Indels 4; Gaps 1;  
 Qy 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNYMHYQKAGQPPKLLIHLASNL 60  
 Db 173 DIELTQSPASLSASVGETVITCRASGNI----HNYLAWYQKQKSPQLLYNNAKTAD 228

Qy 61 GVPARFSGSGCTDFTLNHPVEEDASTYTCQHSGLPFTFGSGTKLEIK 111  
 Db 229 GVPFRSGSGGTQYSLKINSLOPEDFGSYCYQHFWTTPYTFGGGKLEIK 279

RESULT 9  
 Q9ERZ9 PRELIMINARY; PRT; 107 AA.  
 AC Q9ERZ9;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE ANTI HUMAN TNF-ALPHA LIGHT CHAIN VARIABLE REGION (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;  
 RT "Cloning and sequencing of the light chain fragment of variable region  
 RT genes of an anti-hTNF-a monoclonal antibody.";  
 RL J. Cell. Mol. Immunol. 12:21-26(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;  
 RT "Construction and sequencing of the single-chain antibody gene of a  
 RT human TNF-alpha specific monoclonal antibody.";  
 RL Ti 4 Chun i Ta Hsueh Hsueh Pao 19:373-376(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF262753; AAG23804.1; -;  
 FT NON\_TER 107  
 SQ SEQUENCE 107 AA; 11784 MW; 2B15EEA6604A26C3 CRC64;

Query Match 57.4%; Score 332; DB 11; Length 107;  
 Best Local Similarity 59.8%; Pred. No. 4.5e-29;  
 Matches 64; Conservative 18; Mismatches 23; Indels 2; Gaps 1;  
 Qy 4 LTQSPASLAVSLGQRATISCRASKSV--SASGYNYMHYQKAGQPPKLLIHLASNL 61  
 Db 1 MTQSPSSLSAMSVGKVTMSCKSSQSLNSNTQKRYLAWYQKQSPQLLYFASTREG 60

Qy 62 VPARFSGSGCTDFTLNHPVEEDASTYTCQHSGLPFTFGSGTKL 108  
 Db 61 VPDREMGSGCTDFTLTISSVQTEDLADYFCQHQHYTPFTFGSGTKL 107

RESULT 10  
 Q9UL86 PRELIMINARY; PRT; 109 AA.  
 AC Q9UL86;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98277139; PubMed=9614934;  
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,  
 RA Young D.C.;  
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
 RT fetus.";  
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
 CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX  
 CC DOMAIN.

DR EMBL; AF035028; AAD56264.1; -;  
 DR HSP; P01789; IMCP.  
 DR InterPro: IPR003006; -;  
 DR InterPro: IPR003596; -;  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 FT NON\_TER 1  
 FT NON\_TER 109  
 SQ SEQUENCE 109 AA; 11928 MW; 243325F72C7DAC83 CRC64;

Query Match 57.2%; Score 330.5; DB 4; Length 109;  
 Best Local Similarity 59.5%; Pred. No. 6.8e-29;  
 Matches 66; Conservative 18; Mismatches 24; Indels 3; Gaps 1;  
 Qy 1 DIVLTQSPASLAVSLGQRATISCRASKSVASGYNYMHYQKAGQPPKLLIHLASNL 60  
 Db 1 EIVLTQSPGTLSTLFPGERATLSCRASQSVSS--YLAWYQKQPGAPRELLIYGTSSRT 57

Qy 61 GVPARFSGSGCTDFTLNHPVEEDASTYTCQHSGLPFTFGSGTKLEIK 111  
 Db 58 GIPDRFSGSGSETDFTLTISRLPEDEFAVYICQYQYSSITFGPGTKVDIK 108

RESULT 11  
 Q9RIA5 PRELIMINARY; PRT; 214 AA.  
 ID Q9RIA5  
 AC Q9RIA5;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE KAPPA LIGHT CHAIN OF MAB7 (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;  
 RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal  
 RT antibody (Mab 7, its light and heavy chains) and construction of a  
 RT single chain antibody (scFv)."  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF152371; AAD40242.1; -;  
 DR HSP; P01789; IMCP.  
 DR InterPro: IPR003006; -;



Search completed: June 28, 2001, 16:08:26  
Job time: 954 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:14:38 ; Search time 362.28 Seconds  
(without alignments)  
2.510 Million cell updates/sec

Title: US-09-724-406-28

Perfect score: 77

Sequence: 1 RASKSVASGYNVYM 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	94.8	151	22	Antibody variable
2	72	93.5	110	15	Anti-carcinoembryo
3	72	93.5	110	20	Murine COL1 VK cha
4	72	93.5	110	20	Humanised Murine C
5	72	93.5	110	20	Humanised Murine C
6	72	93.5	110	20	Humanised Murine C
7	72	93.5	110	20	Humanised Murine C
8	68	88.3	16	16	Humanised Murine C
9	68	88.3	98	19	L-CDR-1 of anti-id
10	68	88.3	108	12	Light chain of mon
11	68	88.3	110	16	Murine IB4 light c
					Immunoglobulin lig

12	68	88.3	110	16	Immunoglobulin lig
13	68	88.3	110	17	285 light chain va
14	68	88.3	110	18	Monoclonal antibod
15	68	88.3	110	20	Protein sequence o
16	68	88.3	110	20	Protein sequence o
17	68	88.3	110	20	317G5 hybridoma VL
18	68	88.3	110	21	N-terminal sequenc
19	68	88.3	112	11	Light chain variab
20	68	88.3	124	12	HNK-20 variable ka
21	68	88.3	128	17	Anti-idiotypic anti
22	68	88.3	130	16	Anti-idiotypic anti
23	68	88.3	131	13	Anti-idiotypic anti
24	68	88.3	146	16	Mouse anti-idiotyp
25	68	88.3	146	18	Porphylin antibody
26	68	88.3	212	15	H. pylori HSP60-b1
27	64	83.1	15	21	Humanised VL regio
28	64	83.1	111	21	Humanised VL regio
29	61	79.2	111	13	Light chain variab
30	61	79.2	111	15	Murine Fd79 antibo
31	61	79.2	111	22	Humanised Fd79 ant
32	61	79.2	111	22	GM-CSF receptor al
33	60	77.9	22	16	Humanised antibody
34	59	76.6	111	17	Human thyroid stim
35	59	76.6	111	17	CDR1 of the light
36	59	76.6	112	16	Mouse antibody var
37	57	74.0	17	19	Nucleotide sequenc
38	56	72.7	15	16	VL region from an
39	56	72.7	437	19	Light chain variab
40	54	70.1	112	16	Anti-human gp39 MA
41	52	67.5	125	12	Mouse antibody var
42	48	62.3	132	17	Mouse antibody FB3
43	44	57.1	15	16	Mouse antibody F4-
44	44	57.1	218	16	
45	44	57.1	218	16	

## ALIGNMENTS

RESULT 1  
ID AAB35100 standard; Protein; 151 AA.  
AC AAB35100;  
DT 27-MAR-2001 (first entry)  
DE Antibody variable region fusion protein #4.  
KW Filamentous phage; protein display; pVII; pIX;  
KW combinatorial antibody library.  
OS Synthetic.  
PN WO200071694-A1.  
PD 30-NOV-2000.  
PF 24-MAY-2000; 2000WO-US14433.  
PR 25-MAY-1999; 99US-0318786.  
PX (SCRI ) SCRIPPS RES INST.  
PA Janda KD, Wirsching P, Lerner RA, Gao C;  
PI WPI; 2001-032030/04.  
PT Novel filamentous phage encapsulating a genome encoding fusion  
PT polypeptide comprising exogenous polypeptide fused to amino terminus of  
PT pVII and pIX proteins, for constructing diverse heterodimeric  
PT polypeptide array

PS Example 1; Page 82-83; 90pp; English.

XX The present invention describes a filamentous phage encapsulating a  
CC genome encoding a fusion protein. This fusion protein comprises an  
CC exogenous protein fused to the amino terminus of a filamentous phage pVII  
CC or pIX protein. This is useful in the design of proteins for medical,  
CC industrial, environmental and research applications.

SQ Sequence 151 AA;

Query Match 94.8%; Score 73; DB 22; Length 151;  
Best Local Similarity 93.3%; Pred. No. 1.2e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RASKSVSASGYNM 15  
DB 24 rasksvstsgynmh 38  
|||||

RESULT 2

ID AAR60564 standard; Protein; 110 AA.

XX

AC AAR60564;

XX

DT 25-APR-1995 (first entry)

XX

DE Anti-carcinoembryonic antigen chimeric light chain Ab.

XX

KW Anti-carcinoembryonic antigen chimeric antibodies; CEAS;

XX

KW chimeric human-murine; breast or colorectal carcinoma;

XX

XX light chain.

OS Chimeric Mus muscaris.

XX

OS Chimeric Homo sapiens.

XX

PN WO9419466-A.

XX

PD 01-SEP-1994.

XX

PF 16-FEB-1994; 94WO-US01709.

XX

PR 16-FEB-1993; 93US-0017570.

XX

PA (DOWC ) DOW CHEM CO.

XX

PI (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Gourlie BB, Kaplan DA, Mezes PS, Rixon MW, Schlom J;

XX

DR WPI; 1994-294331/36.

XX

DR N-PSDB; AAQ71394.

XX

PT Anti-carcinoembryonic antigen chimeric antibodies - for diagnosis  
PT and therapy of carcinoma, e.g. breast or colorectal carcinoma  
XX

PS Claim 11; Page 50; 67pp; English.

XX

CC AAQ71394 codes for AAR60564 the antibody light chain region of

XX

CC murine-human anti-carcinoembryonic antigen (CEA) chimeric

XX

CC antibody. Which can be used in in vitro immunoassays for

XX

CC the detection of CEA, and monitoring of tumour-associated

XX

CC antigen during therapy. It can also be used in vivo diagnostically,

XX

CC or in therapy for the treatment of tumours associated with

XX

CC colorectal and breast carcinomas, as well those of the

XX

CC gastrointestinal tract, lung, ovary and pancreas.

XX

SQ Sequence 110 AA;

XX

Query Match 93.5%; Score 72; DB 15; Length 110;

XX

Best Local Similarity 93.3%; Pred. No. 1.2e-05;

XX

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX

OY 1 RASKSVSASGYNM 15  
DB 24 rasksvsagsysmh 38  
|||||

RESULT 3

AAAY39532  
ID AAY39532 standard; protein; 110 AA.

XX

AC AAY39532;

XX

DT 22-NOV-1999 (first entry)

XX

DE Murine COL1 VK chain (COL1MuVK).

XX

XX Antibody; human; carcinoembryonic antigen; CEA; cancer; tumour imaging;

KW tumour identification; metastasis; diagnosis; mouse; VK chain.

XX

OS Mus sp.

XX

PN WO9943817-A1.

XX

PD 02-SEP-1999.

XX

PF 25-FEB-1998; 98WO-US03680.

XX

PR 25-FEB-1998; 98WO-US03680.

XX

PA (DOWC ) DOW CHEM CO.

XX

PI Anderson WHK, Tempest PR, Carr FJ, Harris WJ, Armour K;

XX

XX WPI; 1999-550870/46.

XX

XX New antibody specific for treatment and diagnosis of cancer

XX

PS Claim 2; Fig 2; 82pp; English.

XX

CC This sequence represents the murine COL1 VK chain (COL1MuVK), and was  
CC used to obtain the humanised antibody (Ab) of the invention. The  
CC humanised Ab, or its fragment, specifically binds to carcinoembryonic  
CC antigen (CEA). The Ab, optionally coupled to an effector or label, is  
CC used to treat or prevent CEA-expressing cancers (e.g. of breast, ovary,  
CC lung, stomach, or colon) and to detect CEA-expressing cells, either in  
CC vitro (optionally on a solid support) or in vivo (particularly by tumour  
CC imaging to identify tumours and metastases before surgery), for diagnosis  
CC or prognosis. It is not significantly immunogenic, i.e. the Ab does not  
CC induce a human anti-murine antibody or allergic response, or non-specific  
CC cytotoxicity, so can be administered repeatedly. It retains specificity  
CC for CEA, and has improved clearance (allowing efficient targeting) and  
CC metabolic properties.

XX

SQ Sequence 110 AA;

XX

Query Match 93.5%; Score 72; DB 20; Length 110;

XX

Best Local Similarity 93.3%; Pred. No. 1.2e-05;

XX

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RASKSVSASGYNM 15

DB 24 rasksvsagsysmh 38

|||||

RESULT 4

AAAY39534

ID AAY39534 standard; protein; 110 AA.

XX

AC AAY39534;

XX

DT 22-NOV-1999 (first entry)

XX

```

DE Humanised Murine COL1 VK chain (COL1REVK).
XX
KW Antibody; human; carcinoembryonic antigen; CEA; cancer; tumour imaging;
KW tumour identification; metastasis; diagnosis; mouse; VK chain.
XX
OS Synthetic.
OS Mus sp.
OS Homo sapiens.
XX
PN WO9943817-A1.
XX
PD 02-SEP-1999.
XX
PF 25-FEB-1998; 98WO-US03680.
XX
PR 25-FEB-1998; 98WO-US03680.
XX
PX (DOWC ) DOW CHEM CO.
XX
PI Anderson WHK, Tempest PR, Carr FJ, Harris WJ, Armour K;
XX
DR WPI; 1999-550870/46.
XX
PT New antibody specific for treatment and diagnosis of cancer -
XX
PS Claim 2; Fig 2; 82pp; English.
XX
CC This sequence represents a humanised version of the murine COL1 VK chain,
CC designated COL1REVK, and was used to obtain the humanised antibody (Ab)
CC of the invention. The humanised Ab, or its fragment, specifically binds
CC to carcinoembryonic antigen (CEA). The Ab, optionally coupled to an
CC effector or label, is used to treat or prevent CEA-expressing cancers
CC (e.g. of breast, ovary, lung, stomach, or colon) and to detect
CC CEA-expressing cells, either in vitro (optionally on a solid support) or
CC in vivo (particularly by tumour imaging to identify tumours and
CC metastases before surgery), for diagnosis or prognosis. It is not
CC significantly immunogenic, i.e. the Ab does not induce a human
CC anti-murine antibody or allergic response, or non-specific cytotoxicity,
CC so can be administered repeatedly. It retains specificity for CEA, and
CC has improved clearance (allowing efficient targeting) and metabolic
CC properties.
XX
SQ Sequence 110 AA;

Query Match 93.5%; Score 72; DB 20; Length 110;
Best Local Similarity 93.3%; Pred. No. 1.2e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15
Db |||||:||||:||||
24 rasksvsasgysymh 38

RESULT 5
AAY39535
ID AAY39535 standard; protein; 110 AA.
XX
AC AAY39535;
XX
DT 22-NOV-1999 (first entry)
XX
DE Humanised Murine COL1 VK chain (COL1REVK).
XX
KW Antibody; human; carcinoembryonic antigen; CEA; cancer; tumour imaging;
KW tumour identification; metastasis; diagnosis; mouse; VK chain.
XX
OS Synthetic.
OS Mus sp.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 3

```

```

FT Misc-difference 4 /note= "encoded by CAR"
FT FT /note= "encoded by ATG"
FT Misc-difference 32 /note= "encoded by TGT"
XX
PN WO9943817-A1.
XX
PD 02-SEP-1999.
XX
PF 25-FEB-1998; 98WO-US03680.
XX
PR 25-FEB-1998; 98WO-US03680.
XX
PX (DOWC ) DOW CHEM CO.
XX
PI Anderson WHK, Tempest PR, Carr FJ, Harris WJ, Armour K;
XX
DR WPI; 1999-550870/46.
XX
PT N-PSDB: AAZ20530.
XX
PS New antibody specific for treatment and diagnosis of cancer -
XX
PS Claim 2; Fig 14; 82pp; English.
XX
CC This sequence represents a humanised version of the murine COL1 VK chain,
CC designated COL1REVK, and was used to obtain the humanised antibody (Ab)
CC of the invention. The humanised Ab, or its fragment, specifically binds
CC to carcinoembryonic antigen (CEA). The Ab, optionally coupled to an
CC effector or label, is used to treat or prevent CEA-expressing cancers
CC (e.g. of breast, ovary, lung, stomach, or colon) and to detect
CC CEA-expressing cells, either in vitro (optionally on a solid support) or
CC in vivo (particularly by tumour imaging to identify tumours and
CC metastases before surgery), for diagnosis or prognosis. It is not
CC significantly immunogenic, i.e. the Ab does not induce a human
CC anti-murine antibody or allergic response, or non-specific cytotoxicity,
CC so can be administered repeatedly. It retains specificity for CEA, and
CC has improved clearance (allowing efficient targeting) and metabolic
CC properties.
XX
SQ Sequence 110 AA;

Query Match 93.5%; Score 72; DB 20; Length 110;
Best Local Similarity 93.3%; Pred. No. 1.2e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15
Db |||||:||||:||||
24 rasksvsasgysymh 38

RESULT 6
AAY39544
ID AAY39544 standard; protein; 110 AA.
XX
AC AAY39544;
XX
DT 22-NOV-1999 (first entry)
XX
DE Humanised Murine COL1 VK chain variant COL1VK HuVKF.
XX
KW Antibody; human; carcinoembryonic antigen; CEA; cancer; tumour imaging;
KW tumour identification; metastasis; diagnosis; mouse; VK chain.
XX
OS Synthetic.
OS Mus sp.
OS Homo sapiens.
XX
PN WO9943817-A1.
XX
PD 02-SEP-1999.
XX

```

PF 25-FEB-1998; 98WO-US03680.  
 XX  
 PR 25-FEB-1998; 98WO-US03680.  
 XX  
 XX (DOWC ) DOW CHEM CO.  
 PA  
 PI Anderson WHK, Tempest PR, Carr FJ, Harris WJ, Armour K;  
 XX  
 XX WPI; 1999-550870/46.  
 DR  
 XX New antibody specific for treatment and diagnosis of cancer  
 PT  
 PS Disclosure; Page 44; 82pp; English.  
 PS  
 XX This sequence is a humanised variant version of the murine COL1 VK chain,  
 CC designated COL1NMVK, and was used to obtain the humanised antibody (Ab)  
 CC of the invention. The humanised Ab, or its fragment, specifically binds  
 CC to carcinoembryonic antigen (CEA). The Ab, optionally coupled to an  
 CC effector or label, is used to treat or prevent CEA-expressing cancers  
 CC (e.g. of breast, ovary, lung, stomach, or colon) and to detect  
 CC CEA-expressing cells, either in vitro (optionally on a solid support) or  
 CC in vivo (particularly by tumour imaging to identify tumours and  
 CC metastases before surgery), for diagnosis or prognosis. It is not  
 CC significantly immunogenic, i.e. the Ab does not induce a human  
 CC anti-murine antibody or allergic response, or non-specific cytotoxicity,  
 CC so can be administered repeatedly. It retains specificity for CEA, and  
 CC has improved clearance (allowing efficient targeting) and metabolic  
 CC properties.  
 XX Sequence 110 AA;  
 SQ

Query Match 93.5%; Score 72; DB 20; Length 110;  
 Best Local Similarity 93.3%; Pred. No. 1.2e-05;  
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAKSVSASGYNM 15  
 DB 24 rasksvsasgysymh 38  
 |||||:|||||:|||||

RESULT 8  
 AAR74945  
 ID AAR74945 standard; peptide; 16 AA.  
 XX  
 AC AAR74945;  
 XX  
 DT 19-JAN-1996 (first entry)  
 XX  
 DE L-CDR-1 of anti-idiotypic antibody against human anticancer antibody.  
 XX  
 KW Antibody; cancer; CDR; heavy chain; light chain; immunoglobulin;  
 KW complementarity determining region.  
 XX Mus sp.  
 OS  
 XX JP07101999-A.  
 PN  
 XX 18-APR-1995.  
 PD  
 XX 06-OCT-1993; 93JP-0272950.  
 PF  
 XX 06-OCT-1993; 93JP-0272950.  
 PR  
 XX (HAGI/) HAGIWARA Y.  
 PA  
 XX WPI; 1995-182987/24.  
 DR  
 XX Novel anti-idiotypic antibody against a human anticancer monoclonal  
 PT antibody - and DNA sequences encoding the antibody, useful in  
 PT pharmacology, medicine and biochemical fields.  
 XX  
 PS Claim 11; Page 4; 28pp; Japanese.  
 PS  
 XX A new anti-idiotypic antibody against a human anticancer monoclonal  
 CC antibody is claimed. This antibody contains in its heavy chain 3  
 CC complementarity determining regions CDR1 (AAR74929-R74931), CDR2  
 CC (AAR74932-R74935) and CDR3 (AAR74936-R74939), this is also true of the  
 CC light chain which has its own CDR1 (AAR74944-R74946 and AAR85774), CDR2  
 CC (AAR74947-R74949) and CDR3 (AAR74950-R74954). The antibody and DNA  
 CC encoding it are useful in pharmacological, medical and biochemical  
 CC fields.  
 XX Sequence 16 AA;  
 SQ

Query Match 93.5%; Score 72; DB 20; Length 110;  
 Best Local Similarity 93.3%; Pred. No. 1.2e-05;  
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAKSVSASGYNM 15  
 DB 24 rasksvsasgysymh 38  
 |||||:|||||:|||||

RESULT 7  
 AAY39546  
 ID AAY39546 standard; protein; 110 AA.  
 XX  
 AC AAY39546;  
 XX  
 DT 22-NOV-1999 (first entry)  
 XX  
 DE Humanised Murine COL1 VK chain variant.  
 XX  
 KW Antibody; human; carcinoembryonic antigen; CEA; cancer; tumour imaging;  
 KW tumour identification; metastasis; diagnosis; mouse; VK chain.  
 XX  
 OS Synthetic.  
 OS Mus sp.  
 OS Homo sapiens.  
 XX  
 PN WO9943817-A1.  
 PD  
 XX 02-SEP-1999.  
 PF  
 XX 25-FEB-1998; 98WO-US03680.  
 PF  
 XX 25-FEB-1998; 98WO-US03680.  
 PR  
 XX (DOWC ) DOW CHEM CO.  
 PA  
 PI Anderson WHK, Tempest PR, Carr FJ, Harris WJ, Armour K;  
 XX  
 XX WPI; 1999-550870/46.  
 DR  
 XX New antibody specific for treatment and diagnosis of cancer  
 PT

Query Match 88.3%; Score 68; DB 16; Length 16;  
 Best Local Similarity 86.7%; Pred. No. 6.8e-06;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15  
 DB 2 RASKSVSTSGYSYM 16  
 DB 15 RASKSVSTSGYSYM 29

RESULT 9  
 AAW70952  
 ID AAW70952 standard; protein; 98 AA.  
 AC AAW70952;  
 XX  
 XX  
 DT 14-OCT-1998 (first entry)  
 DE Light chain of monoclonal antibody RS-2B8.  
 XX  
 XX Pathogenic virus; tropism; mucosa; CDR region; monoclonal antibody;  
 KW respiratory syncytial virus; RSV; VP6 protein; rota virus; RV;  
 KW viral infection; inhibit; fusion; protection; transcription;  
 KW antiviral agent; prophylaxis; diagnosis; infection; contamination.  
 XX  
 OS Mus sp.  
 XX  
 XX FR2758331-A1.  
 PN  
 XX  
 PD 17-JUL-1998.  
 XX  
 PF 14-JAN-1997; 97FR-0000300.  
 XX  
 PR 14-JAN-1997; 97FR-0000300.  
 XX  
 XX (UYBO-) UNIV BOURGOGNE.  
 PA  
 XX Bourgeois C, Kohl E, Pothier P;  
 PI  
 XX WPI; 1998-390320/34.  
 DR  
 XX  
 PT New peptide(s) recognising viral epitope with tropism to mucosa -  
 PT useful for, e.g. diagnosing, preventing and treating viral  
 PT infection(s)  
 XX  
 PS Disclosure; Fig 2; Sipp; French.  
 CC  
 CC The present sequence represents the light chain of monoclonal antibody  
 CC RS-2B8, which is directed against the respiratory syncytial virus (RSV).  
 CC The specification describes peptides which recognise, by  
 CC antigen-antibody type reactions, at least 1 epitope of a pathogenic virus  
 CC having tropism for the mucosa. AAW70905-16 and AAW70929-46 are analogous  
 CC to CDR regions of monoclonal antibodies specific for RSV. AAW70917-28 are  
 CC analogous to CDR regions of monoclonal antibodies specific for site III  
 CC or IV of the VP6 protein of rota virus (RV). The peptides can neutralise  
 CC viral infections and may also inhibit fusion between infected and  
 CC uninfected cells or cells and viruses. They provide passive or active  
 CC protection and/or inhibit transcription of the virus, so are useful as  
 CC antiviral agents or for prophylaxis, in human or veterinary medicine. The  
 CC peptides can be labelled and used to diagnose infection or contamination  
 CC by the virus. The peptides are particularly directed against RSV or RS  
 CC but may also be used against papilloma, adeno, entero, polio, influenza  
 CC or immune deficiency viruses.  
 XX  
 XX Sequence 98 AA;

Query Match 88.3%; Score 68; DB 19; Length 98;  
 Best Local Similarity 86.7%; Pred. No. 5.6e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15

DB 15 RASKSVSTSGYSYM 29

RESULT 10  
 AAR13088  
 ID AAR13088 standard; Protein; 108 AA.  
 XX  
 AC AAR13088;  
 XX  
 XX 01-OCT-1991 (first entry)  
 DT  
 XX Murine 1B4 light chain-1 variable region.  
 DE  
 XX Monoclonal antibody; complementarity determining region; CDR;  
 KW integrin; hybridoma 1B4; protein REI; Gal/Rei; Ig.  
 KW  
 XX EP438312-A.  
 PN  
 PN EP440351-A.  
 XX  
 XX 24-JUL-1991.  
 PD  
 XX 17-JAN-1991; 91EP-0300367.  
 PF  
 XX 20-DEC-1990; 90US-0627421.  
 PR  
 PR 19-JAN-1990; 90US-0467692.  
 XX  
 XX (MERI ) MERCK & CO INC.  
 PA  
 XX Law MF, Mark GE, Schmidt JA, Singer II;  
 PI  
 XX WPI; 1991-216985/30.  
 DR  
 DR N-PSDB; AAQ12683.  
 XX  
 XX New recombinant immunoglobulin(s) reactive with leukocyte CD18  
 PT antigen - comprise human heavy chain framework and murine  
 PT complementarity regions useful in treatment of inflammation  
 XX  
 PS Disclosure; Fig 24; 77pp; English.  
 CC  
 CC A recombinant human Ig comprises a human heavy chain framework  
 CC and murine CDRs (with the heavy chain framework mutated at sites  
 CC near the CDRs), a human light chain framework and murine CDRs.  
 CC It has a mean IC50 nearly equal to that of the murine monoclonal  
 CC antibody from which the CDRs were derived. It is designated  
 CC mutated Gal/Rei. The human Ig is capable of binding to a human  
 CC CD18 integrin.  
 CC The murine CDRs are obtd. from murine hybridoma 1B4  
 CC (ATCC HB 10164). The light chain framework is derived from  
 CC human myeloma protein REI (EP-239400).  
 CC See also AAQ12682-84.  
 XX  
 SQ Sequence 108 AA;

Query Match 88.3%; Score 68; DB 12; Length 108;  
 Best Local Similarity 86.7%; Pred. No. 6.3e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15  
 DB 24 RASKSVSTSGYSYM 38

RESULT 11  
 AAR74956  
 ID AAR74956 standard; Protein; 110 AA.  
 XX  
 AC AAR74956;  
 XX  
 DT 19-JAN-1996 (first entry)  
 XX  
 XX Immunoglobulin light chain of anti-idiotype antibody against human

DE anticancer antibody.  
 XX Antibody; cancer; CDR: heavy chain; light chain; immunoglobulin;  
 KW complementarity determining region.  
 XX Mus sp.  
 XX OS  
 XX JP07101999-A.  
 XX PN  
 XX PD  
 XX 18-APR-1995.  
 XX PF  
 XX 06-OCT-1993; 93JP-0272950.  
 XX PR  
 XX 06-OCT-1993; 93JP-0272950.  
 XX PA (HAGI/) HAGIWARA Y.  
 XX XX  
 XX WPI; 1995-182987/24.  
 XX DR N-PSDB; AAQ90421.  
 XX XX

XX Novel anti-idiotype antibody against an human anticancer monoclonal  
 PT antibody - and DNA sequences encoding the antibody, useful in  
 PT pharmacology, medicine and biochemical fields.  
 XX PS  
 XX Claim 13; Page 4; 28pp; Japanese.  
 XX XX

CC AAR74955-R74959 are possible light chains of a new anti-idiotype  
 CC antibody against a human anticancer monoclonal antibody. This antibody  
 CC contains in its heavy chain 3 complementarity determining regions CDR1  
 CC (AAR74929-R74931), CDR2 (AAR74932-R74935) and CDR3 (AAR74936-R74939),  
 CC this is also true of the light chain which has its own CDR1  
 CC (AAR74944-R74946 and AAR85774), CDR2 (AAR74947-R74949) and CDR3  
 CC (AAR74950-R74954). The antibody and DNA encoding it are useful in  
 CC pharmacological, medical and biochemical fields.  
 XX XX

SQ Sequence 110 AA;

Query Match 88.3%; Score 68; DB 16; Length 110;  
 Best Local Similarity 86.7%; Pred. No. 6.4e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15  
 ||||| |||:|  
 Db 24 rasksvstgysymh 38

RESULT 12

AAR74957  
 ID AAR74957 standard; Protein; 110 AA.

XX AC AAR74957;

XX DT 22-JAN-1996 (first entry)

XX Immunoglobulin light chain of anti-idiotype antibody against human  
 DE anticancer antibody.

XX Antibody; cancer; CDR: heavy chain; light chain; immunoglobulin;  
 KW complementarity determining region.

XX OS Mus sp.

XX PN JP07101999-A.

XX PD 18-APR-1995.

XX PF 06-OCT-1993; 93JP-0272950.

XX PR 06-OCT-1993; 93JP-0272950.

XX PA (HAGI/) HAGIWARA Y.

XX XX

DR WPI; 1995-182987/24.  
 DR N-PSDB; AAQ90423.  
 XX Novel anti-idiotype antibody against an human anticancer monoclonal  
 PT antibody - and DNA sequences encoding the antibody, useful in  
 PT pharmacology, medicine and biochemical fields.  
 XX PS  
 XX Claim 14; Page 4; 28pp; Japanese.  
 XX XX

CC AAR74955-R74959 are possible light chains of a new anti-idiotype  
 CC antibody against a human anticancer monoclonal antibody. This antibody  
 CC contains in its heavy chain 3 complementarity determining regions CDR1  
 CC (AAR74929-R74931), CDR2 (AAR74932-R74935) and CDR3 (AAR74936-R74939),  
 CC this is also true of the light chain which has its own CDR1  
 CC (AAR74944-R74946 and AAR85774), CDR2 (AAR74947-R74949) and CDR3  
 CC (AAR74950-R74954). The antibody and DNA encoding it are useful in  
 CC pharmacological, medical and biochemical fields.  
 XX XX

SQ Sequence 110 AA;

Query Match 88.3%; Score 68; DB 16; Length 110;  
 Best Local Similarity 86.7%; Pred. No. 6.4e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15  
 ||||| |||:|  
 Db 24 rasksvstgysymh 38

RESULT 13

AAR90832

ID AAR90832 standard; Protein; 110 AA.

XX AC AAR90832;

XX DT 25-JUN-1996 (first entry)

XX DE 2B5 light chain variable region from pCIB4616.

XX delta endotoxin; Bacillus thuringiensis; western corn rootworm;  
 KW WCRW; maize; pesticide; brush border membrane vesicle; monoclonal;  
 KW antibody.

XX OS Insecta sp.

XX PN WO9600783-A1.

XX PD 11-JAN-1996.

XX PF 20-JUN-1995; 95WO-IB00497.

XX PR 28-JUN-1994; 94US-0267641.

XX PA (CIBA ) CIBA GEIGY AG.

XX PI Carozzi NB, Koziel MG;

XX WPI; 1996-077494/08.

XX DR N-PSDB; AAT15728.

XX New monoclonal antibodies which bind insect gut proteins - used  
 PT partic. with toxin moieties for the control of insect pests, partic.  
 PT in plants

XX PS Claim 8; Page 57; 106pp; English.

XX AAR90829-39 are monoclonal antibodies or a binding fragments produced by  
 CC using insect guts, partic. insect brush border membranes (BBMs), esp.  
 CC corn rootworm, as antigen; immunising a donor animal with the antigen;  
 CC isolating immunocompetent B cells from the immunised animal; fusing B  
 CC cells with a tumour cell line; isolating the fused cells, culturing them  
 CC and cloning positive hybrid cells; and screening the hybrid cells for

CC prodn. of the required MABs. The MABs bind to the gut of a target insect  
 CC but do not bind to mammalian BBMs. The DNA sequence can be operably  
 CC linked to a toxin moiety, esp. selected from e.g. Bacillus toxins,  
 CC Pseudomonas exotoxin and phytoalexin, etc.. The Abs are useful for  
 CC control of insect pests, e.g. Coleoptera, Diptera, Hymenoptera and  
 CC Lepidoptera. The pesticidal compsn. is pref. applied to a plant, e.g.  
 CC maize.  
 XX  
 SQ Sequence 110 AA;

Query Match 88.3%; Score 68; DB 17; Length 110;  
 Best Local Similarity 86.7%; Pred. No. 6.4e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15  
 ||||| |||:|  
 Db 24 rasksvstgysmh 38

RESULT 14

AAW22952  
 ID AAW22952 standard; Protein; 110 AA.

XX AC AAW22952;

XX DT 02-APR-1998 (first entry)

XX DE Monoclonal antibody (MAB) 340 light chain variable region.

XX MAB 340; monoclonal antibody; foetal cell; pre-natal diagnosis;  
 KW erythrocyte; trophoblast; genetic abnormality; foetal sex; analysis;  
 KW epidermal growth factor receptor; EGFR; Down's syndrome; human.  
 XX OS Mus sp.

XX FH Key

XX FT Region Location/Qualifiers  
 1..23  
 FT /note= "Framework 1"

FT Region 24..38

FT /note= "Complementarity-determining region 1"

FT Region 39..53

FT /note= "Framework 2"

FT Region 54..60

FT /note= "Complementarity-determining region 2"

FT Region 61..92

FT /note= "Framework 3"

FT Region 93..100

FT /note= "Complementarity-determining region 3"

FT Region 101..110

FT /note= "Framework 4"

XX PN W09730354-A2.

XX PD 21-AUG-1997.

XX PF 17-FEB-1997; 97WO-GB00443.

XX PR 16-FEB-1996; 96GB-0003249.

XX PA (UYNO-) UNIV NOTTINGHAM.

XX XX Durrant LG, Liu DT;

XX PI WPI; 1997-425169/39.

XX DR N-PSDB; AAT75584.

XX XX Foetal cell analysis comprising isolating at least 2 sorts of foetal

PT cell from maternal sample - specifically nucleated red cells and

PT trophoblasts, useful for pre-natal diagnosis of genetic

PT abnormalities or foetal sex

XX XX Example 2; Page 15; 31pp; English.

PS

XX

CC This is the light chain variable region of a monoclonal antibody (MAB)  
 CC 340. The MAB 340 acts as a trophoblast binding agent in a method for  
 CC analysing human foetal cells. The method comprises isolating at least 2  
 CC types of nucleated foetal cell from a maternal sample. The foetal cells  
 CC specifically are nucleated red cells (erythrocytes) and trophoblasts.  
 CC The trophoblasts are isolated by contacting the maternal sample with a  
 CC specific binding agent that binds to the epidermal growth factor receptor  
 CC (EGFR). MAB340 or an EGFR binding derivative of MAB340 comprising at  
 CC least the CDR regions of one or other of the light or heavy chain  
 CC variable regions of MAB 340 is used as the binding agent. Erythrocytes  
 CC are also isolated with a specific anti-transferrin antibody. The target  
 CC nucleotide sequences in the foetal cells are amplified by PCR primers.  
 CC The isolated cells can be used for genetic and biochemical analysis,  
 CC particularly to determine the sex of the foetus or to detect inherited  
 CC abnormalities, e.g. Down's syndrome. More generally MAB 340 can be used  
 CC to enrich (for subsequent analysis) rare malignant cells, e.g. squamous  
 CC carcinoma cells, that express EGF receptor. The method is not  
 CC significantly invasive and does not carry the risk of inducing abortion.  
 CC Use of two types of cells significantly improves the sensitivity of  
 CC pre-natal diagnosis.  
 XX  
 SQ Sequence 110 AA;

Query Match 88.3%; Score 68; DB 18; Length 110;

Best Local Similarity 86.7%; Pred. No. 6.4e-05;

Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15

||||| |||:|

Db 24 rasksvstgysmh 38

RESULT 15

AAW86123

ID AAW86123 standard; Protein; 110 AA.

XX AC AAW86123;

XX DT 03-MAR-1999 (first entry)

XX DE Protein sequence of de-immunised 340 vl.

XX KW Non-immunogenic; epitope; T-cell; immunogenicity; immune system; SK;  
 KW immunoglobulin; therapeutic; streptokinase; de-immunised.  
 XX OS Homo sapiens.  
 XX PN W09852976-A1.  
 XX PD 26-NOV-1998.  
 XX PF 21-MAY-1998; 98WO-GB01473.  
 XX PR 14-APR-1998; 98GB-0007751.  
 XX PR 21-MAY-1997; 97GB-0010480.  
 XX PR 31-JUL-1997; 97GB-0016197.  
 XX PR 28-NOV-1997; 97GB-0025270.  
 XX PR 02-DEC-1997; 97US-0067235.  
 XX XX (BIOV-) BIOVATION LTD.

XX PI Carr FU;

XX DR WPI; 1999-045301/04.

XX XX Reducing immunogenicity of proteins - by modifying the amino acid

PT sequence of the protein to eliminate potential epitopes for T-cells

PT of a given species

XX XX Example 1; Fig 5; 77pp; English.

XX PS

CC The invention relates to a method for the production of non-immunogenic  
 CC proteins. The method comprises determining at least part of the amino  
 CC acid sequence of the protein; (b) identifying in the amino acid sequence  
 CC one or more potential epitopes for T-cells (T-cell epitopes) of the given  
 CC species; and (c) modifying the amino acid sequence to eliminate at least  
 CC one of the T-cell epitopes identified in step (b) thereby to eliminate or  
 CC reduce the immunogenicity of the protein when exposed to the immune  
 CC system of the given species. A method of analysing a pre-existing protein  
 CC to predict the basis for immunogenic responses is also provided. The  
 CC methods can be used particularly for reducing the immunogenicity of  
 CC immunoglobulins or therapeutic proteins, e.g. Streptokinase (SK). The  
 CC products can be used for diagnosis and therapy. The present sequence  
 CC represents the protein sequence of de-immunised 340 Vh.  
 XX  
 SQ Sequence 110 AA;

Query Match 88.3%; Score 68; DB 20; Length 110;  
 Best Local Similarity 86.7%; Pred. No. 6.4e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVASAGYNYMH 15  
 Db 24 rasksvstsgysymh 38

Search completed: June 28, 2001, 16:14:38  
 Job time: 1326 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:01:15 ; Search time 138.34 Seconds  
(without alignments)  
2.184 Million cell updates/sec

Title: US-09-724-406-28  
Perfect score: 77  
Sequence: 1 RASKSVSASGYNYMH 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	93.5	110	1	US-08-017-570-2
2	72	93.5	110	1	US-08-471-426-2
3	72	93.5	110	5	PCT-US94-01709-2
4	69	89.6	20	2	US-08-712-212-1
5	69	89.6	20	5	PCT-US95-05160-1
6	68	88.3	15	1	US-08-137-117D-136
7	68	88.3	15	2	US-08-436-717-136
8	68	88.3	16	1	US-08-318-970B-13
9	68	88.3	110	3	US-08-442-542-8
10	68	88.3	110	3	US-08-483-749A-20
11	68	88.3	110	3	US-08-765-469-8
12	68	88.3	130	2	US-08-137-117D-86
13	68	88.3	130	2	US-08-436-717-86
14	68	88.3	131	1	US-08-137-117D-25
15	68	88.3	131	2	US-08-436-717-25
16	68	88.3	146	2	US-08-653-402B-12
17	61	79.2	111	1	US-07-634-278-46
18	61	79.2	111	1	US-07-634-278-47
19	61	79.2	111	1	US-08-477-728-46
20	61	79.2	111	1	US-08-477-728-47
21	61	79.2	111	1	US-08-207-169A-4
22	61	79.2	111	1	US-08-474-040-46
23	61	79.2	111	1	US-08-474-040-47
24	61	79.2	111	1	US-08-487-200-47
25	61	79.2	111	1	US-08-487-200-46
26	61	79.2	111	4	US-08-484-537-46
27	61	79.2	111	4	US-08-484-537-47

28	60	77.9	128	5	PCT-US95-15716-6	Sequence 6, Appl
29	56	72.7	111	5	PCT-US94-14106-53	Sequence 53, Appl
30	48	62.3	132	2	US-08-379-057-16	Sequence 16, Appl
31	44	57.1	218	5	PCT-US94-14106-57	Sequence 57, Appl
32	42	54.5	111	1	US-08-111-080-30	Sequence 30, Appl
33	42	54.5	111	1	US-08-111-080-32	Sequence 32, Appl
34	42	54.5	111	1	US-08-275-053-11	Sequence 11, Appl
35	42	54.5	111	1	US-08-275-053-14	Sequence 14, Appl
36	42	54.5	111	1	US-08-211-980-30	Sequence 30, Appl
37	42	54.5	111	1	US-08-211-980-32	Sequence 32, Appl
38	42	54.5	111	5	PCT-US93-07967-30	Sequence 30, Appl
39	42	54.5	111	5	PCT-US93-07967-32	Sequence 32, Appl
40	42	54.5	121	1	US-08-111-080-18	Sequence 18, Appl
41	42	54.5	121	1	US-08-111-080-22	Sequence 22, Appl
42	42	54.5	121	1	US-08-211-980-18	Sequence 18, Appl
43	42	54.5	121	1	US-08-211-980-22	Sequence 22, Appl
44	42	54.5	121	5	PCT-US92-07111-17	Sequence 17, Appl
45	42	54.5	121	5	PCT-US93-07967-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-08-017-570-2  
; Sequence 2, Application US/08017570  
; Patent No. 5472693  
; GENERAL INFORMATION:  
; APPLICANT: GOURLIE, BRIAN B  
; APPLICANT: RIXON, MARK W  
; APPLICANT: MEZES, PETER S  
; APPLICANT: KAPLAN, DONALD A  
; APPLICANT: SCHLOM, JEFFREY  
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC  
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Duane C. Ulmer  
; STREET: P.O. Box 1967  
; CITY: Midland  
; STATE: MI  
; COUNTRY: US  
; ZIP: 48641-1967  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 19930216  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ULMER, DUANE C  
; REGISTRATION NUMBER: 34,941  
; REFERENCE/DOCKET NUMBER: C-38,777  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (517) 636-8104  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 110 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-017-570-2

Query Match 93.5%; Score 72: DB 1; Length 110;  
Best Local Similarity 93.3%; Pred. No. 3.5e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RASKSVSASGYNYMH 15  
|||||

```

Db      24 RASKSVSASGYSYMH 38

RESULT 2
US-08-471-426-2
; Sequence 2, Application us/08471426
; Patent No. 5808033
; GENERAL INFORMATION:
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: MEZES, PETER S
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Duane C. Ulmer
; STREET: P.O. Box 1967
; CITY: Midland
; STATE: MI
; COUNTRY: US
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,426
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/017,570
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-38,777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-426-2

Query Match 93.5%; Score 72; DB 1; Length 110;
Best Local Similarity 93.3%; Pred. No. 3.5e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 RASKSVSASGYSYMH 15
        |||
Db      24 RASKSVSASGYSYMH 38

RESULT 3
PCT-US94-01709-2
; Sequence 2, Application PC/TUS9401709
; GENERAL INFORMATION:
; APPLICANT: THE DOW CHEMICAL COMPANY
; APPLICANT: U.S.A. DEPT. OF HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Duane C. Ulmer
; STREET: P.O. Box 1967
; CITY: Midland
; STATE: MI

```

```

; COUNTRY: US
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01709
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: 38,777-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01709-2

Query Match 93.5%; Score 72; DB 5; Length 110;
Best Local Similarity 93.3%; Pred. No. 3.5e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 RASKSVSASGYSYMH 15
        |||
Db      24 RASKSVSASGYSYMH 38

RESULT 4
US-08-712-212-1
; Sequence 1, Application US/08712212
; Patent No. 5837460
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Kieber-Emmons, Thomas
; APPLICANT: Weiner, David B.
; APPLICANT: Vonfeldt, Joan M.
; TITLE OF INVENTION: Biologically active peptides and
; TITLE OF INVENTION: methods of identifying the same
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5837460is
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,212
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,404
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-1554
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100

```

TELEFAX: 215-568-3439  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-712-212-1

Query Match 89.6%; Score 69; DB 5; Length 20;  
 Best Local Similarity 86.7%; Pred. No. 1.7e-05;  
 Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASKSVASGYNMH 15

Db 2 RASKSVSSGYSYMH 16

RESULT 5  
 PCT-US95-05160-1  
 Sequence 1, Application PC/TUS9505160  
 GENERAL INFORMATION:  
 APPLICANT: Williams, William V.  
 APPLICANT: Kieber-Emmons, Thomas  
 APPLICANT: Weiner, David B.  
 APPLICANT: VonFeldt, Joan M.  
 TITLE OF INVENTION: Biologically active peptides and  
 TITLE OF INVENTION: methods of identifying the same  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
 ADDRESSEE: Norris  
 STREET: One Liberty Place, 46th Floor  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Wordperfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/05160  
 FILING DATE:

CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/235,404  
 FILING DATE: 29-APR-1994  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: DeLuca, Mark  
 REGISTRATION NUMBER: 33,229  
 REFERENCE/DOCKET NUMBER: UPN-2245  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-568-3100  
 TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US95-05160-1

Query Match 89.6%; Score 69; DB 5; Length 20;  
 Best Local Similarity 86.7%; Pred. No. 1.7e-05;  
 Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASKSVASGYNMH 15

|||||:|||||

Db 2 RASKSVSSGYSYMH 16

RESULT 6

US-08-137-117D-136  
 Sequence 136, Application US/08137117D  
 Patent No. 5795965  
 GENERAL INFORMATION:

APPLICANT: TSUCHIYA, Masayuki  
 APPLICANT: SATO, Koh  
 APPLICANT: BENDIG, Mary  
 APPLICANT: JONES, Steven  
 APPLICANT: SALJANHA, Jose  
 TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
 TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
 NUMBER OF SEQUENCES: 158  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 3000 K Street, N.W., Suite 500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/137,117D  
 FILING DATE: 20-DEC-1993  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/JP92/00544  
 FILING DATE: 24-APR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 4-32084  
 FILING DATE: 19-FEB-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 3-95476  
 FILING DATE: 25-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WEGNER, Harold C.  
 REGISTRATION NUMBER: 25,258  
 REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)672-5300  
 TELEFAX: (202)672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 136:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 15 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-137-117D-136

Query Match 88.3%; Score 68; DB 1; Length 15;  
 Best Local Similarity 86.7%; Pred. No. 1.9e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVASGYNMH 15

Db 1 RASKSVSTGYSYMH 15

RESULT 7

US-08-436-717-136  
 Sequence 136, Application US/08436717  
 Patent No. 5817790  
 GENERAL INFORMATION:

APPLICANT: TSUCHIYA, Masayuki

APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,717  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117  
FILING DATE: 20-DEC-1993  
APPLICATION NUMBER: WO PCT/JP92/00544  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-32084  
FILING DATE: 19-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-95476  
FILING DATE: 25-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 136:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-436-717-136

Query Match 88.3%; Score 68; DB 2; Length 15;  
Best Local Similarity 86.7%; Pred. No. 1.9e-05;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
DB 1 RASKSVSTSGYSYMH 15

RESULT 8  
US-08-318-970B-13  
Sequence 13, Application US/08318970B  
Patent No. 5589573  
GENERAL INFORMATION:  
APPLICANT: HIDEAKI HAGIWARA, et al.  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC  
TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY  
TITLE OF INVENTION: AND DNA BASE SEQUENCES ENCODING THOSE SEQUENCES  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Office of Sherman and Shalloway  
STREET: 413 N. Washington Street

CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: Dell System 210; Intel 80 285 Microprocessor  
OPERATING SYSTEM: MS DOS 3.3  
SOFTWARE: Word Perfect, Version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,970B  
FILING DATE: October 6, 1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Richard A. Steinberg  
REGISTRATION NUMBER: 26,588  
REFERENCE/DOCKET NUMBER: S-2371  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 549-2282  
TELEFAX: (703) 836-0106  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: L CDRL-2  
OTHER INFORMATION: hypervariable region  
US-08-318-970B-13

Query Match 88.3%; Score 68; DB 1; Length 16;  
Best Local Similarity 86.7%; Pred. No. 2e-05;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
DB 2 RASKSVSTSGYSYMH 16

RESULT 9  
US-08-442-542-8  
Sequence 8, Application US/08442542  
Patent No. 5686600  
GENERAL INFORMATION:  
APPLICANT: Carozzi, Nadine B.  
APPLICANT: Kozziel, Michael G.  
TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
TITLE OF INVENTION: Proteins and their Use  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,542  
FILING DATE: 16-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/267,641  
FILING DATE: 28-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: CGC 1750  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-442-342-8

Query Match 88.3%; Score 68; DB 1; Length 110;  
Best Local Similarity 86.7%; Pred. No. 0.00016;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RASKSVSASGYNM 15  
Db 24 RASKSVSTSGYSYM 38

RESULT 10

US-08-483-749A-20  
Sequence 20, Application US/08483749A  
Patent No. 6054561  
GENERAL INFORMATION:  
APPLICANT: KING, DAVID B.  
TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHIRON CORPORATION  
STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097  
CITY: EMERYVILLE  
STATE: CA  
COUNTRY: USA  
ZIP: 94662-8097  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,749A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: SAVERIDE, PAUL B.  
REGISTRATION NUMBER: 36,914  
REFERENCE/DOCKET NUMBER: 0508.008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2585  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-483-749A-20

Query Match 88.3%; Score 68; DB 3; Length 110;  
Best Local Similarity 86.7%; Pred. No. 0.00016;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RASKSVSASGYNM 15  
Db 24 RASKSVSTSGYSYM 38

RESULT 11

US-08-765-469-8  
Sequence 8, Application US/08765469  
Patent No. 6069301  
GENERAL INFORMATION:  
APPLICANT: Carozzi, Nadine B.  
APPLICANT: Kozziel, Michael G.  
TITLE OF INVENTION: Antibodies which Bind to Insect Gut  
TITLE OF INVENTION: Proteins and their Use  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/765,469  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/267,641  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1750  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-765-469-8

Query Match 88.3%; Score 68; DB 3; Length 110;  
Best Local Similarity 86.7%; Pred. No. 0.00016;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RASKSVSASGYNM 15  
Db 24 RASKSVSTSGYSYM 38

RESULT 12

US-08-137-117D-86  
Sequence 86, Application US/08137117D  
Patent No. 5795965  
GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, Masayuki  
APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117D  
FILING DATE: 20-DEC-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP92/00544  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-32084  
FILING DATE: 19-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-95476  
FILING DATE: 25-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53456/126/AAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 130 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-137-117D-86

Query Match 88.3%; Score 68; DB 1; Length 130;  
Best Local Similarity 86.7%; Pred. No. 0.00019;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15  
||||||| |||:||||

DB 43 RASKSVSTSGYSYM 57

RESULT 13  
US-08-436-717-86  
Sequence 86, Application US/08436717  
Patent No. 5817790  
GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, Masayuki  
APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,717  
FILING DATE:  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117  
FILING DATE: 20-DEC-1993  
APPLICATION NUMBER: WO PCT/JP92/00544  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-32084  
FILING DATE: 19-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-95476  
FILING DATE: 25-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 130 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-436-717-86

Query Match 88.3%; Score 68; DB 2; Length 130;  
Best Local Similarity 86.7%; Pred. No. 0.00019;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNM 15

DB 43 RASKSVSTSGYSYM 57

RESULT 14  
US-08-137-117D-25  
Sequence 25, Application US/08137117D  
Patent No. 5795965  
GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, Masayuki  
APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117D  
FILING DATE: 20-DEC-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP92/00544  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-32084  
FILING DATE: 19-FEB-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 3-95476  
 FILING DATE: 25-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WEGNER, Harold C.  
 REGISTRATION NUMBER: 25,258  
 REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)672-5300  
 TELEFAX: (202)672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 131 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-137-117D-25

Query Match 88.3%; Score 68; DB 1; Length 131;  
 Best Local Similarity 86.7%; Pred. No. 0.00019;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
 ||||| |||:  
 Db 44 RASKSVSTSGYSYMH 58

RESULT 15

US-08-436-717-25  
 Sequence 25, Application US/08436717  
 Patent No. 5817790

GENERAL INFORMATION:

APPLICANT: TSUCHIYA, Masayuki  
 APPLICANT: SATO, Koh  
 APPLICANT: BENDIG, Mary  
 APPLICANT: JONES, Steven  
 APPLICANT: SALDANHA, Jose  
 TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
 TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
 NUMBER OF SEQUENCES: 158  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 3000 K Street, N.W., Suite 500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/436.717  
 FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/137.117

FILING DATE: 20-DEC-1993  
 APPLICATION NUMBER: WO PCT/JP92/00544  
 FILING DATE: 24-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 4-32084  
 FILING DATE: 19-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 3-95476  
 FILING DATE: 25-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: WEGNER, Harold C.  
 REGISTRATION NUMBER: 25,258  
 REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300  
 TELEFAX: (202)672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 131 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-436-717-25

Query Match 88.3%; Score 68; DB 2; Length 131;  
 Best Local Similarity 86.7%; Pred. No. 0.00019;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
 ||||| |||:  
 Db 44 RASKSVSTSGYSYMH 58

Search completed: June 28, 2001, 16:01:15  
 Job time: 523 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:58:46 ; Search time 234.85 Seconds  
(without alignments)  
4.865 Million cell updates/sec

Title: US-09-724-406-28  
Perfect score: 77  
Sequence: 1 RASKSVSASGYNM 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	89.6	210	2 A56169	Ig kappa chain v r
2	68	88.3	102	2 PC6027	acetylcoline recep
3	68	88.3	108	2 PH0092	Ig kappa chain v r
4	68	88.3	110	2 S24288	Ig kappa chain v r
5	68	88.3	111	1 KVM584	Ig kappa chain v r
6	68	88.3	111	1 KVM575	Ig kappa chain v r
7	68	88.3	111	1 KVM585	Ig kappa chain v r
8	68	88.3	112	2 S45715	Ig kappa chain v r
9	68	88.3	128	2 S52448	Ig kappa chain v r
10	68	88.3	128	2 JL0073	aberrant kappa tra
11	68	88.3	140	2 PN0446	Ig kappa chain v r
12	68	88.3	218	2 S68241	Ig kappa chain v r
13	68	88.3	218	2 JC5810	monoclonal antibod
14	68	88.3	233	2 JC5322	p53 specific singl
15	66	85.7	111	1 KVM5A0	Ig kappa chain v r
16	65	84.4	111	2 S09563	Ig kappa chain v r
17	64	83.1	108	1 KVM5S4	Ig kappa chain v r
18	59	76.6	131	2 S55027	Ig light chain v r
19	58	75.3	102	2 PH1076	Ig light chain v r
20	54	70.1	102	2 PH1077	Ig light chain v r
21	53	68.8	101	2 S59640	Ig light chain v r
22	52	67.5	115	2 S63596	Ig kappa chain v r
23	43	55.8	109	2 PH0093	Ig kappa chain v r
24	42	54.5	91	2 S25462	Ig kappa chain v r
25	42	54.5	107	2 S26343	Ig kappa chain v r
26	42	54.5	107	2 S26344	Ig kappa chain v r
27	42	54.5	111	1 KVM537	Ig kappa chain v r
28	42	54.5	111	2 D45722	anti-glycoprotein
29	42	54.5	131	1 KVM5M6	Ig kappa chain pre

30 42 54.5 764 2 T48446  
31 41 53.2 111 2 E53285  
32 41 53.2 112 2 S19972  
33 41 53.2 221 2 A56941  
34 41 53.2 233 2 T01233  
35 41 53.2 301 2 B61164  
36 41 53.2 608 2 B53295  
37 41 53.2 2693 2 A40743  
38 40 51.9 320 1 A39479  
39 40 51.9 321 1 S28390  
40 40 51.9 346 2 T20458  
41 40 51.9 662 2 B75544  
42 39 50.6 112 2 S51122  
43 38.5 50.0 94 2 PL0258  
44 38 49.4 96 2 S17608  
45 38 49.4 96 2 B49442

ALIGNMENTS

RESULT 1  
A56169  
Ig kappa chain V region (clone 23.2) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 11-Jan-2000  
C;Accession: A56169  
R;Monfardini, C.; Kieber-Emmons, T.; VonFeldt, J.M.; O'Malley, B.; Rosenbaum, H.; God  
J. Biol. Chem. 270, 6628-6638, 1995  
A;Title: Recombinant antibodies in bioactive peptide design.  
A;Reference number: A56169; MUID:95204454  
A;Accession: A56169  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-210 <MON>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotrimer; immunoglobulin

Query Match 89.6%; Score 69; DB 2; Length 210;  
Best Local Similarity 86.7%; Pred. No. 0.00013;  
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RASKSVSASGYNM 15  
DB 24 RASKSVSSSGSYMH 38  
RESULT 2  
PC6027  
acetylcoline receptor monoclonal antibody A7 light chain variable region - mouse (fra  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 21-Jan-2000  
C;Accession: PC6027  
R;Meng, F.; Yang, K.; Graus, Y.; de Baets, M.  
Chinese J. Microbiol. Immunol. 16, 45-48, 1996  
A;Title: Pathogenicity and sequence analysis of a mouse anti-AChR monoclonal antibody  
A;Reference number: PC6026  
A;Accession: PC6027  
A;Molecule type: mRNA  
A;Residues: 1-102 <MEN>  
A;Cross-references: EMBL:X80961  
C;Comment: This protein induces acetylcoline receptor loss and muscular weakness in m  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
F;8-86/Domain: immunoglobulin homology <IMM>

Query Match 88.3%; Score 68; DB 2; Length 102;  
Best Local Similarity 86.7%; Pred. No. 9.1e-05;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RASKSVSASGYNM 15  
||||||| |||||

Db 16 RASKSVSTSGSYMH 30

RESULT 3

Ig kappa chain V region (anti-cyclosporin F) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 21-Jan-2000  
 C:Accession: PH0092  
 R:Schmitter, D.; Poch, O.; Zeder, G.; Heinrich, G.F.; Kocher, H.P.; Quesniaux, V.F.J.; V  
 Mol. Immunol. 27, 1029-1038, 1990  
 A:Title: Analysis of the structural diversity of monoclonal antibodies to cyclosporine.  
 A:Reference number: PH0087; MUID:91042649  
 A:Accession: PH0092  
 A:Molecule type: mRNA  
 A:Residues: 1-108 <SCH>  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:24-38/Region: complementarity-determining 1  
 F:54-60/Region: complementarity-determining 2  
 F:93-100/Region: complementarity-determining 3

Query Match 88.3%; Score 68; DB 2; Length 108;  
 Best Local Similarity 86.7%; Pred. No. 9.6e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RASKSVSASGYNMH 15  
 ||||| |||:|  
 Db 24 RASKSVSTSGSYMH 38

RESULT 4

Ig kappa chain V region (JS28/32) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000  
 C:Accession: S24288  
 R:Moncharmont, B.  
 submitted to the EMBL Data Library, September 1991  
 A:Description: Cloning and sequencing of the cDNA coding for the variable regions of the  
 A:Reference number: S24287  
 A:Accession: S24288  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-110 <MON>  
 A:Cross-references: EMBL:X62703; NID:g51673; PIDN:CAA44576.1; PID:g1333958  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:16-94/Domain: immunoglobulin homology <IMM>

Query Match 88.3%; Score 68; DB 2; Length 110;  
 Best Local Similarity 86.7%; Pred. No. 9.8e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RASKSVSASGYNMH 15  
 ||||| |||:|  
 Db 24 RASKSVSTSGSYMH 38

RESULT 5

Ig kappa chain V region (PC6684) - mouse (tentative sequence)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 31-Mar-2000  
 C:Accession: A01938  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: A01938

A:Molecule type: protein  
 A:Residues: 1-111 <WEI>  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 88.3%; Score 68; DB 1; Length 111;  
 Best Local Similarity 86.7%; Pred. No. 9.9e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RASKSVSASGYNMH 15  
 ||||| |||:|  
 Db 24 RASKSVSTSGSYMH 38

RESULT 6

Ig kappa chain V region (PC7175) - mouse (tentative sequence)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 31-Mar-2000  
 C:Accession: B01938; A01938  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: B01938  
 A:Molecule type: protein  
 A:Residues: 1-111 <WEI>  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 88.3%; Score 68; DB 1; Length 111;  
 Best Local Similarity 86.7%; Pred. No. 9.9e-05;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RASKSVSASGYNMH 15  
 ||||| |||:|  
 Db 24 RASKSVSTSGSYMH 38

RESULT 7

Ig kappa chain V regions (PC2485, PC4039) - mouse (tentative sequence)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 31-Mar-2000  
 C:Accession: A01939  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: A01939  
 A:Molecule type: protein  
 A:Residues: 1-111 <WEI>  
 A:Note: the PC4285 and PC4039 sequences are identical  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 88.3%; Score 68; DB 1; Length 111;  
 Best Local Similarity 86.7%; Pred. No. 9.9e-05;

Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
||||||| |||:|

Db 24 RASKSVSTSGYSYMH 38

# RESULT 8

S45715  
Ig kappa chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 19-Mar-1997 #sequence\_revision 14-Nov-1997 #text\_change 07-May-1999

C:Accession: S45715

R:Kim, H.H.; Kato, K.; Yamato, S.; Igarashi, T.; Matsunaga, C.; Ohtsuka, H.; Higuchi, A.  
FEBS Lett. 346, 246-250, 1994

A:Title: Application of (13)C NMR spectroscopy to paratope mapping for larger antigen-Peptide  
A:Reference number: S45714; MUID:94283606

A:Accession: S45715

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-112 <KIM>

A:Experimental source: cell line Ex-3C7

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:16-94/Domain: immunoglobulin homology <IMM>

F:23-92/Disulfide bonds: #status predicted

Query Match 88.3%; Score 68; DB 2; Length 112;  
Best Local Similarity 86.7%; Pred. No. 0.0001;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
||||||| |||:|

Db 24 RASKSVSTSGYSYMH 38

# RESULT 9

S52448

Ig kappa chain V region - mouse

C:Species: Mus musculus (house mouse)

C>Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 21-Jan-2000

C:Accession: S52448

R:Berdoz, J.; Kraehenbuhl, J.P.

submitted to the EMBL Data Library, November 1994

A:Description: Specific amplification by the polymerase chain reaction of rearranged ger

A:Reference number: S52445

A:Accession: S52445

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-128 <BER>

A:Cross-references: EMBL:X82689

C:Genetics:

A:Introns: 14/1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:37-115/Domain: immunoglobulin homology <IMM>

Query Match 88.3%; Score 68; DB 2; Length 128;  
Best Local Similarity 86.7%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
||||||| |||:|

Db 45 RASKSVSTSGYSYMH 59

# RESULT 10

JL0073

aberrant kappa transcript - mouse

C:Species: Mus musculus (house mouse)

C>Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 21-Jan-2000

C:Accession: JL0073

R:Carroll, W.L.; Mendel, E.; Levy, S.

Mol. Immunol. 25, 991-995, 1988

A:Title: Hybridoma fusion cell lines contain an aberrant kappa transcript.

A:Reference number: JL0073; MUID:89112230

A:Accession: JL0073

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-128 <CAR>

A:Cross-references: GB:M35669; NID:g197295; PIDN:AAA39004.1; PID:g309369

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:36-114/Domain: immunoglobulin homology <IMM>

Query Match 88.3%; Score 68; DB 2; Length 128;  
Best Local Similarity 86.7%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
||||||| |||:|

Db 44 RASKSVSTSGYSYMH 58

# RESULT 11

PN0446

Ig kappa chain precursor V-II region - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000

C:Accession: PN0446

R:Kaluzza, R.; Betzl, G.; Shao, H.; Diamantstein, T.; Weidle, U.H.

Gene 122, 321-328, 1992

A:Title: A general method for chimerization of monoclonal antibodies by inverse polym

A:Reference number: PN0444; MUID:93138402

A:Accession: PN0446

A:Molecule type: mRNA

A:Residues: 1-140 <KAL>

A:Cross-references: GB:L02345

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-140/Product: Ig light chain kappa-2 V region #status predicted <MAT>

F:36-114/Domain: immunoglobulin homology <IMM>

Query Match 88.3%; Score 68; DB 2; Length 140;  
Best Local Similarity 86.7%; Pred. No. 0.00013;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
||||||| |||:|

Db 44 RASKSVSTSGYSYMH 58

# RESULT 12

S68241

Ig kappa chain V region (Mab13-1) - mouse (fragment)

N:Alternate names: immunoglobulin light chain

C:Species: Mus musculus (house mouse)

C>Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Jun-2000

C:Accession: S68241; S68214

R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka,

submitted to the EMBL Data Library, March 1994

A:Description: Specific peroxidase activity by formation of an antibody L-chain-porph

A:Reference number: S68241

A:Accession: S68241

A:Molecule type: mRNA

A:Residues: 1-218 <TAK>

A:Cross-references: EMBL:D29670; NID:g473962; PIDN:BAA06141.1; PID:g473963

R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka,

FEBS Lett. 375, 273-276, 1995

A:Title: Thermolabile peroxidase activity with a recombinant antibody L chain-porphyr

A:Reference number: S68211; MUID:96085223

A:Accession: S68214

A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 'NI', 3-212 <TAW>  
A:Cross-references: EMBL:D29670  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin

Query Match 88.3%; Score 68; DB 2; Length 218;  
Best Local Similarity 93.3%; Pred. No. 0.0002;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RASKSVSASGYNYMH 15  
DB 24 RASKSVSASGYIYMH 38  
|||||  
|||||

RESULT 13  
JC5810  
monoclonal antibody 13-1 light chain - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jan-2000  
C:Accession: JC5810  
R:Akashi, S.; Kato, K.; Torizawa, T.; Dohmae, N.; Yamaguchi, H.; Kamachi, M.; Harada, A.  
Biochem. Biophys. Res. Commun. 240, 566-572, 1997  
A:Title: Structural characterization of mouse monoclonal antibody 13-1 against a porphyrin  
A:Reference number: JC5810; MUID:98063277  
A:Accession: JC5810  
A:Molecule type: protein  
A:Residues: 1-218 <AKA>  
C:Comment: This catalytic antibody has peroxidase oxidase. It is directed against a porphyrin  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:16-94/Domain: immunoglobulin homology <IMM>

Query Match 88.3%; Score 68; DB 2; Length 218;  
Best Local Similarity 93.3%; Pred. No. 0.0002;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RASKSVSASGYNYMH 15  
DB 24 RASKSVSASGYIYMH 38  
|||||  
|||||

RESULT 14  
JC5322  
p53 specific single-chain antibody Pab421 - human  
C:Species: Homo sapiens (man)  
C:Date: 15-May-1997 #sequence\_revision 15-May-1997 #text\_change 18-Jul-1997  
C:Accession: JC5322  
R:Janot, C.B.; Hynes, N.E.  
Biochem. Biophys. Res. Commun. 230, 242-246, 1997  
A:Title: Characterization of scFv-421, a single-chain antibody targeted to p53.  
A:Reference number: JC5322; MUID:97168950  
A:Accession: JC5322  
A:Molecule type: mRNA  
A:Residues: 1-233 <JAN>  
A:Experimental source: hydrioloma cell  
C:Comment: This protein specifically binds the tumor suppressor protein p53. It restores

Query Match 88.3%; Score 68; DB 2; Length 233;  
Best Local Similarity 86.7%; Pred. No. 0.00021;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RASKSVSASGYNYMH 15  
DB 148 RASKSVSTSGYSYMH 162  
|||||  
|||||

RESULT 15  
KWS40  
Ig kappa chain V region (PC7940) - mouse (tentative sequence)

C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 31-Mar-2000  
C:Accession: C01938; A01938  
R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
Nature 276, 785-790, 1978  
A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
A:Reference number: A93204; MUID:79073152  
A:Accession: C01938  
A:Molecule type: protein  
A:Residues: 1-111 <WEI>  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:16-94/Domain: immunoglobulin homology <IMM>  
F:23-92/Disulfide bonds: #status predicted

Query Match 85.7%; Score 66; DB 1; Length 111;  
Best Local Similarity 86.7%; Pred. No. 0.00022;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RASKSVSASGYNYMH 15  
DB 24 RASKSVSASGYIYMH 38  
|||||  
|||||

Search completed: June 28, 2001, 15:58:46  
Job time: 374 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:54:39 ; Search time 105.36 Seconds  
(without alignments)  
4.877 Million cell updates/sec

Title: US-09-724-406-28

Perfect score: 77

Sequence: 1 RASKSVASGYNYMH 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	88.3	111	1 KV3R_MOUSE	P01670 mus musculus
2	68	88.3	111	1 KV3S_MOUSE	P01671 mus musculus
3	68	88.3	111	1 KV3T_MOUSE	P01673 mus musculus
4	66	85.7	111	1 KV3U_MOUSE	P01672 mus musculus
5	64	83.1	108	1 KV3V_MOUSE	P01674 mus musculus
6	42	54.5	111	1 KV3W_MOUSE	P01660 mus musculus
7	42	54.5	111	1 KV3J_MOUSE	P01662 mus musculus
8	42	54.5	111	1 KV3K_MOUSE	P01663 mus musculus
9	42	54.5	131	1 KV3I_MOUSE	P01661 mus musculus
10	41	53.2	221	1 LEC6_TACTR	P82151 tachypleus
11	41	53.2	301	1 YGBI_HAEIN	P45979 haemophilus
12	41	53.2	608	1 XNCC_FIBSU	P35811 fibrobacter
13	40	51.9	320	1 MEC3_CAEVU	P34765 caenorhabdi
14	40	51.9	321	1 MEC3_CAEEL	P09088 caenorhabdi
15	40	51.9	2468	1 MAPB_HUMAN	P46821 homo sapien
16	39	50.6	111	1 KV3C_MOUSE	P01656 mus musculus
17	38	49.4	132	1 KV3F_MOUSE	P01658 mus musculus
18	38	49.4	334	1 ARGC_BUCAL	P57156 buchnera ap
19	38	49.4	816	1 NP42_MOUSE	P97460 mus musculus
20	37.5	48.7	421	1 P2X5_HUMAN	Q93086 homo sapien
21	37	48.1	265	1 KKH1_CANGA	O74684 candida gla
22	37	48.1	407	1 CHLP_SYNY3	Q55087 synechocyst
23	36.5	47.4	374	1 PEL3_ERWCA	P29171 erwinia car
24	36.5	47.4	375	1 PELC_ERWCH	P11073 erwinia chr
25	36	46.8	111	1 KV3L_MOUSE	P01664 mus musculus
26	36	46.8	239	1 US21_HCMVA	P09723 human cytom
27	36	46.8	233	1 YE24_MYCTU	P71693 mycobacteri
28	36	46.8	457	1 SGCC_ECOLI	P39365 escherichia
29	36	46.8	590	1 YW72_YEAST	Q05021 saccharomyc
30	35.5	46.1	1651	1 CO3_NAJNA	Q01833 najia najia (
31	35	45.5	121	1 RL20_CHLPN	Q926r7 chlamydia p
32	35	45.5	123	1 RL20_CHLTR	O84842 chlamydia t
33	35	45.5	195	1 YW72_INBLE	P03493 influenza b

34 35 45.5 295 1 MEC3\_CAEBR P34764 caenorhabdi  
35 35 45.5 378 1 AGAL\_COFAR Q42656 coffea arab  
36 35 45.5 419 1 CLPX\_CHLTR O84711 chlamydia t  
37 35 45.5 445 1 6PGD\_CITDI P41582 citrobacter  
38 35 45.5 445 1 6PGD\_CITER P41583 citrobacter  
39 35 45.5 445 1 6PGD\_ESCVU P41574 escherichia  
40 35 45.5 445 1 6PGD\_KLEPL P41575 klebsiella  
41 35 45.5 445 1 6PGD\_KLETE P41577 klebsiella  
42 35 45.5 445 1 6PGD\_SHIBO P41578 shigella bo  
43 35 45.5 445 1 6PGD\_SHIDY P41579 shigella dy  
44 35 45.5 445 1 6PGD\_SHISO P41580 shigella so  
45 35 45.5 460 1 HEMN\_PSEAE P77915 pseudomonas

#### ALIGNMENTS

RESULT 1  
KV3R\_MOUSE  
ID KV3R\_MOUSE STANDARD; PRT; 111 AA.  
AC P01670;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG KAPPA CHAIN V-III REGION PC 6684.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=79073152; PubMed=103003;  
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
RT "Rearrangement of genetic information may produce immunoglobulin  
RT diversity.";  
RL Nature 276:785-790(1978).  
DR PIR: A01938: KVM84;  
DR InterPro: IPR003006;  
DR Pfam: PF00047; Ig: 1.  
KW Immunoglobulin V region.  
FT DOMAIN 1 23 FRAMEWORK 1.  
FT DOMAIN 24 38 COMPLEMENTARITY-DETERMINING 1.  
FT DOMAIN 39 53 FRAMEWORK 2.  
FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING 2.  
FT DOMAIN 61 92 FRAMEWORK 3.  
FT DOMAIN 93 101 COMPLEMENTARITY-DETERMINING 3.  
FT DOMAIN 102 111 FRAMEWORK 4.  
FT DISULFID 23 92 BY SIMILARITY.  
FT NON\_TER 111 111  
SQ SEQUENCE 111 AA; 12039 MW; 1E46988341858526 CRC64;

Query Match 88.3%; Score 68; DB 1; Length 111;  
Best Local Similarity 86.7%; Pred No. 3e-05;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVASGYNYMH 15  
DB 24 RASKSVSTSGSYMH 38  
||||| |||:|

RESULT 2  
KV3S\_MOUSE  
ID KV3S\_MOUSE STANDARD; PRT; 111 AA.  
AC P01671;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG KAPPA CHAIN V-III REGION PC 7175.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

```

RN SEQUENCE.
RP MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
RT diversity.";
RL Nature 276:785-790(1978).
DR PIR; B01938; KWS75.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12010 MW; F041E89AA7858523 CRC64;

Query Match 88.3%; Score 68; DB 1; Length 111;
Best Local Similarity 86.7%; Pred. No. 3e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15
DB 24 RASKSVSTSGYSYMH 38

RESULT 3
KV3U_MOUSE STANDARD; PRT; 111 AA.
AC P01673;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG KAPPA CHAIN V-III REGION PC 2485/PC 4039.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
RT diversity.";
RL Nature 276:785-790(1978).
DR PIR; A01939; KWS85.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 11986 MW; BF38C59AA7858467 CRC64;

Query Match 88.3%; Score 68; DB 1; Length 111;
Best Local Similarity 86.7%; Pred. No. 3e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15

```

```

DB 24 RASKSVSTSGYSYMH 38

RESULT 4
KV3T_MOUSE STANDARD; PRT; 111 AA.
AC P01672;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG KAPPA CHAIN V-III REGION PC 7940.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
RT diversity.";
RL Nature 276:785-790(1978).
DR PIR; C01938; KWS40.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12038 MW; EE8A82306084352E CRC64;

Query Match 85.7%; Score 66; DB 1; Length 111;
Best Local Similarity 86.7%; Pred. No. 6.8e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15
DB 24 RASKSVSAFGYSYMH 38

RESULT 5
KV3V_MOUSE STANDARD; PRT; 108 AA.
AC P01674;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG KAPPA CHAIN V-III REGION PC 2154.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
RT diversity.";
RL Nature 276:785-790(1978).
DR PIR; A01940; KWS54.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38

```



FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 108  
 FT DISULFID 23 92  
 FT NON\_TER 108 108  
 SQ SEQUENCE 108 AA; 11699 MW; D40921D18DAC4B9E CRC64;

Query Match 83.1%; Score 64; DB 1; Length 108;  
 Best Local Similarity 80.0%; Pred. No. 0.00015;  
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RASKSVSASGYNYMH 15  
 |||:|||||:|||||  
 DB 24 RASQSVSTSGYSYMH 38

RESULT 6  
 KV3H\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01660;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 3741/TEPC 111.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE (PC 3741);  
 RX MEDLINE=79073152; PubMed=99744;  
 RA McKean D.J., Bell M., Potter M.;  
 RT "Mechanisms of antibody diversity: multiple genes encode structurally  
 RT related mouse kappa variable regions."  
 RT Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).  
 RL Nature 276:785-790(1978).  
 RN [2]

FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12099 MW; EC46C9D259213BE4 CRC64;

Query Match 54.5%; Score 42; DB 1; Length 111;  
 Best Local Similarity 53.3%; Pred. No. 1.2;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 RASKSVSASGYNYMH 15  
 |||:|||||:|||||  
 DB 24 RASESVDSYNSFMH 38

RESULT 7  
 KV3J\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01663;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 4050.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 RT diversity."  
 RT Nature 276:785-790(1978).  
 RL PIR; A01935; KVM5M6.  
 DR InterPro; IPR003006; -.  
 DR Pfam; PF00047; Ig; 1.  
 DR Immunoglobulin V region.  
 KW Immunoglobulin V region.

Query Match 54.5%; Score 42; DB 1; Length 111;  
 Best Local Similarity 53.3%; Pred. No. 1.2;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 RASKSVSASGYNYMH 15  
 |||:|||||:|||||  
 DB 24 RASESVDSYNSFMH 38

RESULT 7  
 KV3J\_MOUSE

ID KV3J\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01662;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION ABPC 22/PC 9245.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE (ABPC 22).  
 RX MEDLINE=79012520; PubMed=99744;  
 RA McKean D.J., Bell M., Potter M.;  
 RT "Mechanisms of antibody diversity: multiple genes encode structurally  
 RT related mouse kappa variable regions."  
 RT Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).  
 RL Nature 276:785-790(1978).  
 RN [2]  
 RP SEQUENCE (PC 9245);  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 RT diversity."  
 RT Nature 276:785-790(1978).  
 CC -I- MISCELLANEOUS: THE ABPC22 AND PC9241 SEQUENCES ARE IDENTICAL.  
 DR PIR; A01935; KVM5M6.  
 DR InterPro; IPR003006; -.  
 DR Pfam; PF00047; Ig; 1.  
 DR Immunoglobulin V region.  
 KW Immunoglobulin V region.  
 FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12041 MW; D7DF06099303453CE CRC64;

Query Match 54.5%; Score 42; DB 1; Length 111;  
 Best Local Similarity 53.3%; Pred. No. 1.2;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 RASKSVSASGYNYMH 15  
 |||:|||||:|||||  
 DB 24 RASESVDSYNSFMH 38

RESULT 8  
 KV3K\_MOUSE STANDARD; PRT; 111 AA.  
 AC P01663;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION PC 4050.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=79073152; PubMed=103003;  
 RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
 RT "Rearrangement of genetic information may produce immunoglobulin  
 RT diversity."  
 RT Nature 276:785-790(1978).  
 RL PIR; A01935; KVM5M6.  
 DR InterPro; IPR003006; -.  
 DR Pfam; PF00047; Ig; 1.  
 DR Immunoglobulin V region.

FT DOMAIN 1 23  
 FT DOMAIN 24 38  
 FT DOMAIN 39 53  
 FT DOMAIN 54 60  
 FT DOMAIN 61 92  
 FT DOMAIN 93 101  
 FT DOMAIN 102 111  
 FT DISULFID 23 92  
 FT NON\_TER 111 111  
 SQ SEQUENCE 111 AA; 12005 MW; 39D87619313453CB CRC64;

Query Match 54.5%; Score 42; DB 1; Length 111;  
 Best Local Similarity 53.3%; Pred. No. 1.2;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 RASKSVASGYNYMH 15  
 |||:|:|:|:|:  
 DB 24 RASESVDSYGNFSMH 38

RESULT 9  
 KV31\_MOUSE STANDARD; PRT; 131 AA.  
 AC P01661;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG KAPPA CHAIN V-III REGION MOPC 63 PRECURSOR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE OF 1-35.  
 RX MEDLINE=78235887; PubMed=98179;  
 RA Burstein Y., Schechter I.;  
 RT "Primary structures of N-terminal extra peptide segments linked to  
 the variable and constant regions of immunoglobulin light chain  
 precursors: implications on the organization and controlled  
 expression of immunoglobulin genes.";  
 RL Biochemistry 17:2392-2400(1978).  
 RN [2]

RP SEQUENCE OF 21-131.  
 RX MEDLINE=73140225; PubMed=4691517;  
 RA McKean D.J., Potter M., Hood L.E.;  
 RT "Mouse immunoglobulin chains. Pattern of sequence variation among  
 kappa chains with limited sequence differences.";  
 RL Biochemistry 12:760-771(1973).  
 RN [3]

RP REVISIONS.  
 RX MEDLINE=79012520; PubMed=99744;  
 RA McKean D.J., Bell M., Potter M.;  
 RT "Mechanisms of antibody diversity: multiple genes encode structurally  
 related mouse kappa variable regions.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).

DR PIR: A01935; KVM5M6;  
 DR InterPro: IPR003006;  
 DR Pfam: PF00047; Ig; 1.

KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 131  
 FT DOMAIN 21 131 IG KAPPA CHAIN V-III REGION MOPC 63.  
 FT DOMAIN 21 43  
 FT DOMAIN 44 58  
 FT DOMAIN 59 73  
 FT DOMAIN 74 80  
 FT DOMAIN 81 112  
 FT DOMAIN 113 121  
 FT DOMAIN 122 131  
 FT DISULFID 43 112  
 FT NON\_TER 131 131  
 SQ SEQUENCE 131 AA; 14291 MW; D212EC9F08DC880A CRC64;

Query Match 54.5%; Score 42; DB 1; Length 131;  
 Best Local Similarity 53.3%; Pred. No. 1.4;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 RASKSVASGYNYMH 15  
 |||:|:|:|:|:  
 DB 44 RASESVDSYGNFSMH 58

RESULT 10  
 LEC6\_TACTR STANDARD; PRT; 221 AA.  
 ID LEC6\_TACTR  
 AC P82151;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE LECTIN L6.  
 OS Tachyplesus tridentatus (Japanese horseshoe crab).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
 OC Limulidae; Tachyplesus.  
 OX NCBI\_TaxID=6853;  
 RN [1]

RP SEQUENCE.  
 RC TISSUE=Hemocyte;  
 RX MEDLINE=95301536; PubMed=7782311;  
 RA Saito T., Kawabate S.-I., Hirata M., Iwanaga S.;  
 RT "A novel type of limulus lectin-L6. Purification, primary structure,  
 and antibacterial activity.";  
 RL J. Biol. Chem. 270:14493-14499(1995).  
 CC -!- FUNCTION: LIPOPOLYSACCHARIDE-BINDING PROTEIN WITH GRAM-NEGATIVE  
 ANTIBACTERIAL ACTIVITY. BINDS ZINC AND CALCIUM.

CC -!- SUBCELLULAR LOCATION: L-GRANULES.  
 CC -!- TISSUE SPECIFICITY: HEMOCYTES.  
 CC -!- SIMILARITY: BELONGS TO THE TECTONIN FAMILY.  
 KW Lectin; Antibiotic; Zinc; Calcium; Metal-binding; Alkylation; Repeat.

FT DOMAIN 1 221 6 X APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 1 38 1.  
 FT REPEAT 39 75 2.  
 FT REPEAT 76 113 3.  
 FT REPEAT 114 150 4.  
 FT REPEAT 151 188 5.  
 FT REPEAT 189 221 6.  
 FT MOD\_RES 201 201 ALKYLATION.  
 FT DISULFID 32 36  
 FT DISULFID 108 112  
 FT DISULFID 183 187  
 SQ SEQUENCE 221 AA; 24410 MW; 59C97CF736DF4962 CRC64;

Query Match 53.2%; Score 41; DB 1; Length 221;  
 Best Local Similarity 72.7%; Pred. No. 3.8;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 KSVASGYNYM 14  
 | | | | | | | |  
 DB 86 KHVSAAGYGYI 96

RESULT 11  
 YGBJ\_HAEIN STANDARD; PRT; 301 AA.  
 ID YGBJ\_HAEIN  
 AC P44979;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL PROTEIN HI1010.  
 GN HI1010.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]

SEQUENCE FROM N.A.  
 RP STRAIN-RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Smith K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- SIMILARITY: BELONGS TO THE 3-HYDROXYISOBUTYRATE DEHYDROGENASE  
 CC FAMILY. STRONG, TO E.COLI YGBJ.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U32782; AAC22671.1; -;  
 DR TIGR; H11010; -;  
 DR InterPro: IPR002204; -;  
 DR PROSITE; PS00895; 3-HYDROXYISOBUT\_DH; 1.  
 KW Hypothetical protein; Oxidoreductase; NAD.  
 FT ACT\_SITE 177 177 BY SIMILARITY.  
 SQ SEQUENCE 301 AA; 31035 MW; 7C615A25B0947D18 CRC64;

Query Match 53.2%; Score 41; DB 1; Length 301;  
 Best Local Similarity 46.2%; Pred. NO. 5.2;  
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 SKSVASAGYNYMH 15  
 DB 49 AKAVANGYDFAH 61  
 :.:.:.:.:.::

RESULT 12  
 XINC.FIBSU STANDARD; PRT; 608 AA.  
 ID XINC.FIBSU STANDARD; PRT; 608 AA.  
 AC P35811;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ENDO-1,4-BETA-XYLANASE C PRECURSOR (EC 3.2.1.8) (XYLANASE C)  
 DE (1,4-BETA-D-XYLAN XYLANOXYDROLASE C).  
 GN XINC.  
 OS Fibrobacter succinogenes (Bacteroides succinogenes).  
 OC Bacteria; Fibrobacter/Acidobacteria group; Fibrobacter group;  
 OC Fibrobacter.  
 OX NCBI\_TaxID=833;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-41.  
 RC STRAIN=S85;  
 RX MEDLINE=94064561; PubMed=8244936;  
 RA Paradis F.W., Zhu H., Krell P.J., Phillips J.P., Forsberg C.W.;  
 RT "The xinc gene from Fibrobacter succinogenes S85 codes for a xylanase  
 RT with two similar catalytic domains.";  
 RL J. Bacteriol. 175:7666-7672(1993).  
 CC -1- FUNCTION: CELAVES XYLANS WITH THE PRODUCTION OF XULOSE, XYLOBIOS  
 CC AND XYLO-OLIGOSACCHARIDES.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC  
 CC LINKAGES IN XYLANS.  
 CC -1- PATHWAY: XYLAN DEGRADATION.  
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY G (FAMILY 11 OF GLYCOSYL  
 CC HYDROLASES).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U01037; AAA21848.1; -;  
 DR HSSP; P36218; LXYN  
 DR InterPro: IPR001137; -;  
 DR Pfam; PF00457; Glyco\_Hydro\_11; 2.  
 DR PRINTS; PR00911; GLHYDRLASE1.  
 DR PROSITE; PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
 DR PROSITE; PS00777; GLYCOSYL\_HYDROL\_F11\_2; 1.  
 KW Cellulose degradation; Xylan degradation; Hydrolase; Glycosidase;  
 KW Signal; Repeat. 1  
 FT SIGNAL 25  
 FT CHAIN 26 608 ENDO-1,4-BETA-XYLANASE C.  
 FT DOMAIN 26 262 XYLANASE DOMAIN 1.  
 FT DOMAIN 263 285 SER-RICH (LINKER).  
 FT DOMAIN 286 529 XYLANASE DOMAIN 2.  
 FT DOMAIN 530 543 SER-RICH (LINKER).  
 FT ACT\_SITE 142 142 NUCLEOPHILE (BY SIMILARITY).  
 FT ACT\_SITE 237 237 PROTON DONOR (BY SIMILARITY).  
 FT ACT\_SITE 409 409 NUCLEOPHILE (BY SIMILARITY).  
 FT ACT\_SITE 501 501 PROTON DONOR (BY SIMILARITY).  
 SQ SEQUENCE 608 AA; 66415 MW; 7FA338016F5857CE CRC64;

Query Match 53.2%; Score 41; DB 1; Length 608;  
 Best Local Similarity 66.7%; Pred. NO. 11;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 SKSVASAGYNYMH 14  
 DB 386 SKOGSAGGYNYI 397  
 :.:.:.:.:.::

RESULT 13  
 MEC3.CAEVU STANDARD; PRT; 320 AA.  
 ID MEC3.CAEVU STANDARD; PRT; 320 AA.  
 AC F34765;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE MECHANOSENSORY PROTEIN 3.  
 GN MEC-3.  
 OS Caenorhabditis vulgaris.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=31233;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92084094; PubMed=1684166;  
 RA Way J.C., Wang L., Run J.Q., Wang A.;  
 RT "The mec-3 gene contains cis-acting elements mediating positive and  
 RT negative regulation in cells produced by asymmetric cell division in  
 RT Caenorhabditis elegans.";  
 RL Genes Dev. 5:2193-2211(1991).  
 CC -1- FUNCTION: MEC-3 SPECIFIES DIFFERENTIATION OF THE SET OF SIX TOUCH  
 CC RECEPTOR NEURONS. BINDS COOPERATIVELY AS A HETERODIMER WITH UNC-86  
 CC TO SITES IN THE MEC-3 GENE PROMOTER.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SIMILARITY: TO OTHER HOMEOBOX DOMAINS. BELONGS TO THE LIM  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC -----

DR EMBL: X63956; CAA45377.1; -  
 DR PIR: A39479; A39479.  
 DR HSP: P06601; IFJL.  
 DR TRANSFAC: T01970; -  
 DR InterPro: IPR001356; -  
 DR InterPro: IPR001781; -  
 DR Pfam: PF00412; LIM; 2.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; FALSE\_NEG.  
 DR PROSITE: PS00071; HOMEBOX\_2; 1.  
 DR PROSITE: PS00478; LIM\_DOMAIN\_1; 2.  
 DR PROSITE: PS00023; LIM\_DOMAIN\_2; 2.  
 DR Homeobox; DNA-binding; Nuclear protein; Developmental protein;  
 KW Repeat; LIM motif; Metal-binding; Zinc.  
 KW DOMAIN 29 79 LIM.  
 FT DOMAIN 89 145 HOMEBOX.  
 FT DNA\_BIND 216 275 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 314 320  
 SQ SEQUENCE 320 AA; 36849 MW; AF98BD31C2B2FDB6 CRC64;

Query Match 51.9%; Score 40; DB 1; Length 320;  
 Best Local Similarity 46.7%; Pred. No. 8.3;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 RASKSVSASGNYMH 15  
 : | | | | | | |  
 Db 186 QVKKEVDAYGNFEH 200

RESULT 14  
 MEC3\_CAEEL  
 ID MEC3\_CAEEL STANDARD; PRT; 321 AA.  
 AC P09088;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE MECHANOSENSORY PROTEIN 3.  
 GN MEC-3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88253425; PubMed=2898300;  
 RA Way J.C., Chalfie M.;  
 RT "mec-3, a homeobox-containing gene that specifies differentiation of  
 RT the touch receptor neurons in *C. elegans*.";  
 RL Cell 54:5-16(1988).  
 RN [2]  
 RP REVISIONS. SEQUENCE FROM N.A.  
 RX MEDLINE=93099872; PubMed=1361171;  
 RA Xue D., Finney M., Ruvkun G., Chalfie M.;  
 RT "Regulation of the mec-3 gene by the *C.elegans* homeoproteins UNC-86  
 RT and MEC-3.";  
 RL EMBO J. 11:4969-4979(1992).  
 CC -1- FUNCTION: MEC-3 SPECIFIES DIFFERENTIATION OF THE SET OF SIX TOUCH  
 CC RECEPTOR NEURONS. BINDS COOPERATIVELY AS A HETERODIMER WITH UNC-86  
 CC TO SITES IN THE MEC-3 GENE PROMOTER.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SIMILARITY: TO OTHER HOMEBOX DOMAINS. BELONGS TO THE LIM  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC -----

DR EMBL: L02877; AAA50614.1; -  
 DR EMBL: M20244; AAA28108.1; ALT\_SEQ.  
 DR PIR: A27662; A27662.  
 DR PIR: S28390; S28390.  
 DR HSP: P06601; IFJL.  
 DR TRANSFAC: T01076; -  
 DR InterPro: IPR001356; -  
 DR InterPro: IPR001781; -  
 DR Pfam: PF00412; LIM; 2.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 DR PROSITE: PS00478; LIM\_DOMAIN\_1; 2.  
 DR PROSITE: PS00023; LIM\_DOMAIN\_2; 2.  
 DR PROSITE: PS00071; HOMEBOX\_2; 1.  
 DR Homeobox; DNA-binding; Nuclear protein; Developmental protein;  
 KW Repeat; LIM motif; Metal-binding; Zinc.  
 KW DOMAIN 29 79 LIM.  
 FT DOMAIN 89 145 HOMEBOX.  
 FT DNA\_BIND 217 276 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 315 321  
 SQ SEQUENCE 321 AA; 37088 MW; 9231E7C3F5054853 CRC64;

Query Match 51.9%; Score 40; DB 1; Length 321;  
 Best Local Similarity 46.7%; Pred. No. 8.3;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 RASKSVSASGNYMH 15  
 : | | | | | | |  
 Db 187 QVKKEVDAYGNFEH 201

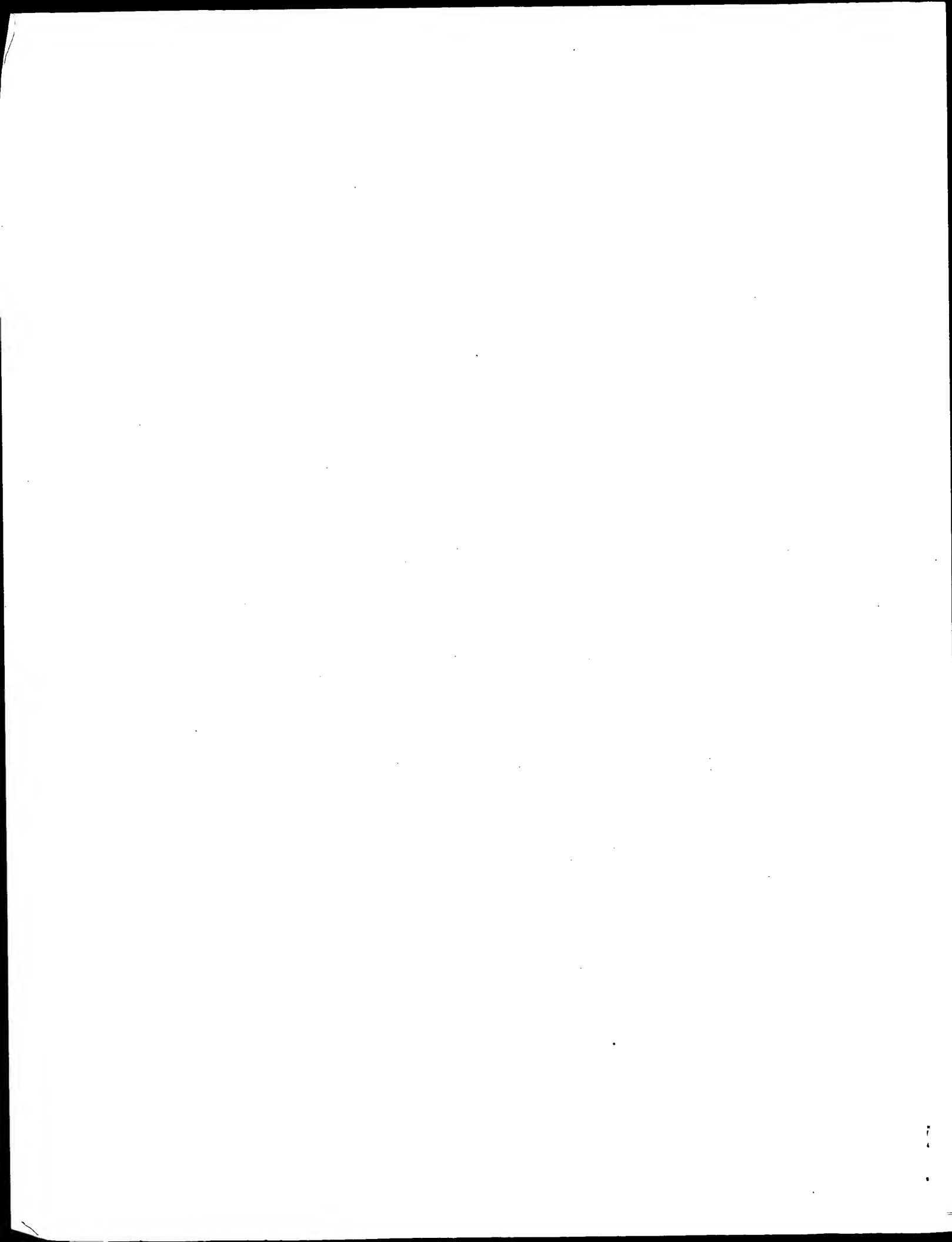
RESULT 15  
 MAPB\_HUMAN  
 ID MAPB\_HUMAN STANDARD; PRT; 2468 AA.  
 AC P46821;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE MICROTUBULE-ASSOCIATED PROTEIN 1B [CONTAINS: MAP1 LIGHT CHAIN LC1].  
 GN MAP1B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95104835; PubMed=7806212;  
 RA Lien L.L., Feener C., Fischbach N., Kunkel L.M.;  
 RT "Cloning of human microtubule-associated protein 1B and the  
 RT identification of a related gene on chromosome 15.";  
 RL Genomics 22:273-280(1994).  
 CC -1- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.  
 CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES  
 CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST  
 CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS  
 CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN  
 CC STABILIZING MICROTUBULES.  
 CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE  
 CC WITH MAP1A AND MAP1B PROTEINS.  
 CC -1- DOMAIN: HAS A HIGHLY BASIC REGION WITH MANY COPIES OF THE SEQUENCE  
 CC KKEE AND KKEI/V, REPEATED BUT NOT AT FIXED INTERVALS. THIS LATTER  
 CC REGION IS RESPONSIBLE FOR THE BINDING OF MAP1B TO MICROTUBULES  
 CC BOTH IN VITRO AND IN VIVO.  
 CC -1- PTM: LC1 IS COEXRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED  
 CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH  
 CC BOTH MAP1A AND MAP1B.

CC -!- SIMILARITY: TO NEURAXIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L06237; AAA18904.1; -  
 DR MIM; 157129; -  
 DR InterPro; IPR000102; -  
 DR Pfam; PF00414; MAP1B\_neuraxin; 10.  
 DR PROSITE; PS00230; MAP1B\_NEURAXIN; 6.  
 KW Microtubules; Repeat; Phosphorylation.  
 FT CHAIN ? 2468 MAP1 LIGHT CHAIN LC1.  
 FT DOMAIN 589 790 LYS-RICH (HIGHLY BASIC, CONTAINS MANY  
 FT 1869 2074 KKEE AND KKEI/V REPEATS).  
 FT REPEAT 1869 1885 12 X 17 AA TANDEM REPEATS.  
 FT REPEAT 1886 1902 1.  
 FT REPEAT 1903 1919 2.  
 FT REPEAT 1920 1936 3.  
 FT REPEAT 1937 1953 4.  
 FT REPEAT 1954 1970 5.  
 FT REPEAT 1971 1987 6.  
 FT REPEAT 1988 2004 7.  
 FT REPEAT 2005 2021 8.  
 FT REPEAT 2022 2038 9.  
 FT REPEAT 2039 2055 10.  
 FT REPEAT 2056 2072 11.  
 FT REPEAT 2073 2090 12.  
 SQ SEQUENCE 2468 AA; 270618 MW; 540839CBDF09D461 CRC64;

Query Match 51.9%; Score 40; DB 1; Length 2468;  
 Best Local Similarity 61.5%; Pred. No. 70;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Oy 1 RASKSVSASGNY 13  
 | : | | | | |  
 Db 1911 RTTKSPSDSGSY 1923

Search completed: June 28, 2001, 15:54:39  
 Job time: 127 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:08:26 ; Search time 411.58 Seconds  
(without alignments)  
4.822 Million cell updates/sec

Title: US-09-724-406-28

Perfect score: 77

Sequence: 1 RASKSVSASGYNYMH 15

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL16:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.rodent:\*
- 12: sp.unclassified:\*
- 13: sp.vertebrate:\*
- 14: sp.virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	57.1	477	10	Q9ZWP7
2	43	55.8	90	2	Q9L817
3	43	55.8	240	2	Q47197
4	43	55.8	240	2	O30925
5	42	54.5	764	10	Q9L275
6	41	53.2	114	3	P78570
7	41	53.2	203	5	Q9U5E9
8	41	53.2	232	5	Q9NXX8
9	41	53.2	233	10	O65278
10	41	53.2	256	5	Q9NB63
11	41	53.2	377	10	Q9FWU1
12	41	53.2	2693	13	Q9L908
13	41	53.2	3469	5	Q9U412
14	41	53.2	3502	5	Q9VYJ9
15	40	51.9	328	5	O02240
16	40	51.9	492	5	O76263
17	40	51.9	662	2	Q9RXX4
18	40	51.9	1049	5	Q9VSV5
19	40	51.9	1095	5	Q9TVG7

20	39.5	51.3	257	2	Q9KJT6	Q9KJT6 staphylococ
21	39	50.6	112	14	Q82092	Q82092 human rhino
22	39	50.6	170	11	Q9QYLO	Q9QYLO mus musculo
23	39	50.6	445	2	Q59473	Q59473 k 6-phospho
24	39	50.6	1122	13	P70047	P70047 xenopus lae
25	39	50.6	1518	5	Q9VN58	Q9VN58 drosophila
26	38	49.4	60	5	Q9NE91	Q9NE91 leishmania
27	38	49.4	150	10	Q9M4Y8	Q9M4Y8 cucumis sat
28	38	49.4	229	2	Q9KYG2	Q9KYG2 streptomyce
29	38	49.4	241	2	Q9KSQ5	Q9KSQ5 vibrio chol
30	38	49.4	254	3	Q9P734	Q9P734 neurospora
31	38	49.4	284	2	Q9WXL1	Q9WXL1 rhodovulum
32	38	49.4	288	10	Q96256	Q96256 arabidopsis
33	38	49.4	292	6	Q9XT99	Q9XT99 oryctolagus
34	38	49.4	317	10	O48720	O48720 arabidopsis
35	38	49.4	377	10	Q9M1P0	Q9M1P0 arabidopsis
36	38	49.4	415	10	Q9SJH8	Q9SJH8 arabidopsis
37	38	49.4	469	5	Q9VF93	Q9VF93 drosophila
38	38	49.4	494	5	Q9NIN7	Q9NIN7 drosophila
39	37.5	48.7	263	2	O54487	O54487 staphylococ
40	37.5	48.7	353	2	Q04086	Q04086 erwinia car
41	37	48.1	56	2	Q9X678	Q9X678 prevotella
42	37	48.1	113	10	O42025	O42025 arabidopsis
43	37	48.1	181	9	Q37970	Q37970 bacterioph
44	37	48.1	245	2	Q9WJ6	Q9WJ6 blastomonas
45	37	48.1	253	2	O32402	O32402 rhodocyclu

#### ALIGNMENTS

RESULT 1

Q9ZWP7 ID Q9ZWP7 PRELIMINARY: PRT: 477 AA.  
AC Q9ZWP7;  
DT 01-MAY-1999 (TREMREL. 10, Created)  
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)  
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)  
DE RNA-DEPENDENT RNA POLYMERASE.  
OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; rosids I;  
OC Rosales; Rosaceae; Pyrus.  
OX NCBI\_TaxID=3767;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99138178; PubMed=9972231;  
RA Osaki H., Kudo A., Ohtsu Y.;  
RT "Nucleotide sequence of seed- and pollen-transmitted double-stranded  
RT RNA, which encodes a putative RNA polymerase, detected  
RT from Japanese pear."  
RL Biosci. Biotechnol. Biochem. 62:2101-2106(1998).  
DR EMBL; AB012616; BAA34783.1; -;  
KW RNA-directed RNA polymerase.  
SQ SEQUENCE 477 AA; 54908 MW; 071E9797E6B1C2F2 CRC64;

Query Match 57.1%; Score 44; DB 10; Length 477;  
Best Local Similarity 72.7%; Pred. No. 8.4;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 KSVSASGYNYM 14

DB 125 KSSTAAGNYNM 135

RESULT 2

Q9L817 ID Q9L817 PRELIMINARY: PRT: 90 AA.  
AC Q9L817;  
DT 01-OCT-2000 (TREMREL. 15, Created)  
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)  
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)

```

DE PUTATIVE G-FIMBRIAL CHAPERONE (FRAGMENT).
GN GAFB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanakanen J.M., Saarela S., Westerlund-Wikstrom B.A., Korhonen T.K.;
RT "The G-fimbrial chaperone protein.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217229; AAF32303.1; -.
DR InterPro; IPR001829; -.
DR Pfam; PF00345; pili_assembly; 1.
DR PRINTS; PR00969; CHAPERONPILI.
DR PROSITE; PS00635; PILI_CHAPERONE; 1.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 10345 MW; 7E1EB1242E0946AA CRC64;

Query Match 55.8%; Score 43; DB 2; Length 90;
Best Local Similarity 61.5%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ASKSVSASGYNYM 14
DB 50 AGKSEKAQGYNYL 62
IIIIII

RESULT 3
Q47197
ID Q47197 PRELIMINARY; PRT; 240 AA.
AC Q47197;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DR EMBL; L77091; AAA92619.1; -.
DR HSP; P31697; 1BF8.
DR InterPro; IPR001829; -.
DR Pfam; PF00345; pili_assembly; 1.
DR PRINTS; PR00969; CHAPERONPILI.
DR PRODOM; PD001447; -.
DR PROSITE; PS00635; PILI_CHAPERONE; 1.
SQ SEQUENCE 240 AA; 26469 MW; C82E5790CF5B9F35 CRC64;

Query Match 55.8%; Score 43; DB 2; Length 240;
Best Local Similarity 61.5%; Pred. No. 6;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ASKSVSASGYNYM 14
DB 119 AKSEKAQGYNYL 131
IIIIII

RESULT 4
O30925
ID O30925 PRELIMINARY; PRT; 240 AA.
AC O30925;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)

```

```

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE F17A-D.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE OF 1-65 FROM N.A.
RA Lintermans P., Pohl P., Deboeck F., Bertels A.M., Schlicker C.,
RT "Isolation and nucleotide sequence of the F17-A gene encoding the
RT structural protein of the F17 fimbriae in bovine enterotoxigenic
RT Escherichia coli.";
RL Infect. Immun. 56:1475-1484(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-25KH09ST;
RX MEDLINE-88226920; PubMed-2897333;
RA Lintermans P., Pohl P., Deboeck F., Bertels A.M., Schlicker C.,
RA Vandekerckhove J.S., Van Damme J., Van Montagu M.M., De Greve H.;
RT "Isolation and nucleotide sequence of the F17-A gene encoding the
RT structural protein of the F17 fimbriae in bovine enterotoxigenic
RT Escherichia coli.";
RL Infect. Immun. 56:1475-1484(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-25KH09ST;
RA De Greve H., Heng P., Deboeck F., Lin H.Y., Schlicker C.,
RA Lintermans P., Hernalsteens J.-P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022140; AAC45720.1; -.
DR HSP; P31697; 1BF8.
DR InterPro; IPR001829; -.
DR Pfam; PF00345; pili_assembly; 1.
DR PRINTS; PR00969; CHAPERONPILI.
DR PRODOM; PD001447; -.
DR PROSITE; PS00635; PILI_CHAPERONE; 1.
SQ SEQUENCE 240 AA; 26455 MW; C83BB790CF5B9F35 CRC64;

Query Match 55.8%; Score 43; DB 2; Length 240;
Best Local Similarity 61.5%; Pred. No. 6;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ASKSVSASGYNYM 14
DB 119 AKSEKAQGYNYL 131
IIIIII

RESULT 5
Q9LZ75
ID Q9LZ75 PRELIMINARY; PRT; 764 AA.
AC Q9LZ75;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 86.3 KDA PROTEIN.
GN T32M21.100.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buysschaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarroel R., Gieken J., Van Montagu M., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162875; CAB85556.1; -.
KW Hypothetical protein.
SQ SEQUENCE 764 AA; 86299 MW; AFE0E482FF091CB8 CRC64;

Query Match 54.5%; Score 42; DB 10; Length 764;

```



Best Local Similarity 50.0%; Pred. No. 33;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASKVSASGYNYMH 15  
      : : : : : : : : : : : : : : : : :  
Db 13 SOKNVAASGHNNHH 26

RESULT 6  
P78570  
ID P78570 PRELIMINARY; PRT; 114 AA.  
AC P78570;  
DT 01-MAY-1997 (TReMBLrel. 03, Created)  
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE DNA-DIRECTED RNA POLYMERASE I (EC 2.7.7.6) (RNA POLYMERASE II) (RNA  
DE POLYMERASE III) (RNA NUCLEOTIDYLTRANSFERASE (DNA-DIRECTED))  
DE (FRAGMENT).  
GN RPAB.  
OS Agaricus bisporus (Common mushroom).  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;  
OC Agaricaceae; Agaricus.  
OX NCBI\_TaxID=5341;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HORST U1;  
RX MEDLINE=97111990; PubMed=8953726;  
RA Sonnenberg A.S.M., De Groot P.W.J., Schaap P.J., Baars J.J.P.,  
Vlaser J., van Griensven L.J.L.D.;  
RT "Isolation of expressed sequence tags of Agaricus bisporus and their  
RT assignment to chromosomes.";  
RL Appl. Environ. Microbiol. 62:4542-4547(1996).  
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE +  
CC RNA(N) (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
DR EMBL; X94765; CAA64391.1; -;  
DR InterPro; IPR001572; -;  
DR Pfam; PF00552; RNA\_pol\_B; 1.  
KW DNA-directed RNA polymerase; Nucleotidyltransferase; Transcription;  
KW Transferase.  
FT NON\_TER 1 1  
FT 114 114  
SQ SEQUENCE 114 AA; 12903 MW; 656C52B0A6C2908 CRC64;

Query Match 53.2%; Score 41; DB 3; Length 114;  
Best Local Similarity 70.0%; Pred. No. 6.1;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 KVSASGYNY 13  
      : : : : : : : : : : : : : : : : :  
Db 20 KQLSAAGYNY 29

RESULT 7  
Q9U5E9  
ID Q9U5E9 PRELIMINARY; PRT; 203 AA.  
AC Q9U5E9;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE TACHYLECTIN-P (FRAGMENT).  
GN TL-P.  
OS Tachypleus tridentatus (Japanese horseshoe crab).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Tachypleus.  
OX NCBI\_TaxID=6853;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20076430; PubMed=10608824;  
RA Nagai T., Kawabata, S., Shishikura F., Sugita H.;  
RT "Purification, characterization, and amino acid sequence of an  
RT embryonic lectin in perivitelline fluid of the horseshoe crab.";

RL J. Biol. Chem. 274:37673-37678(1999).  
DR EMBL; AB028144; BAA88574.1; -;  
FT NON\_TER 1 1  
FT NON\_TER 203 203  
SQ SEQUENCE 203 AA; 22550 MW; 30900733AC2C89D2 CRC64;

Query Match 53.2%; Score 41; DB 5; Length 203;  
Best Local Similarity 72.7%; Pred. No. 11;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 KVSASGYNYN 14  
      : : : : : : : : : : : : : : : : :  
Db 85 KHSASGYGYI 95

RESULT 8  
Q9NKX8  
ID Q9NKX8 PRELIMINARY; PRT; 232 AA.  
AC Q9NKX8;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE TACHYLECTIN-1 EMBRYO (FRAGMENT).  
GN TL-LEMB  
OS Tachypleus tridentatus (Japanese horseshoe crab).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Tachypleus.  
OX NCBI\_TaxID=6853;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=WHOLE BODY;  
RA Nagai T., Kawabata S., Sugita H.;  
RT "Tachylectin-1emb.";  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB037394; BAA90326.1; -;  
FT NON\_TER 1 1  
SQ SEQUENCE 232 AA; 25820 MW; CBD4BFF289F852D1 CRC64;

Query Match 53.2%; Score 41; DB 5; Length 232;  
Best Local Similarity 72.7%; Pred. No. 13;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 KVSASGYNYM 14  
      : : : : : : : : : : : : : : : : :  
Db 85 KHSASGYGYI 95

RESULT 9  
O65278  
ID O65278 PRELIMINARY; PRT; 233 AA.  
AC O65278;  
DT 01-AUG-1998 (TReMBLrel. 07, Created)  
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE F6N23.7 PROTEIN.  
GN F6N23.7 OR AT4G00690.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA WASHU;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Geisel C.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Waterston R.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
[4]  
RN SEQUENCE FROM N.A.  
RP Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
RN SEQUENCE FROM N.A.  
RP EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF058919; AAC13629.1; -;  
DR EMBL: AL161472; CAB80878.1; -;  
DR MEROFS; C48.001; -;  
DR Mendel; 28980; Arath;3300;28980.  
DR InterPro; IPR003653; -;  
KW Hypothetical protein.  
SQ SEQUENCE 233 AA; 27632 MW; 46A1E31F70EC86B8 CRC64;

Query Match 53.2%; Score 41; DB 10; Length 233;  
Best Local Similarity 80.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 KVSASGYN 13  
| | | | |  
Db 85 KLVSGSYN 94

RESULT 10  
Q9NB63 PRELIMINARY; PRT; 256 AA.  
AC Q9NB63;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE GALACTOSE-BINDING PROTEIN  
OS Tachypleus tridentatus (Japanese horseshoe crab).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Tachypleus.  
OX NCBI\_TaxID=6853;  
[1]  
RN SEQUENCE FROM N.A.  
RA Chen S.-C., Yeh M.-S., Chiou S.-T., Huang C.-J., Liu T.-Y.;  
RT "Genomic structure of galactose-binding protein from the hemolymph of  
RT Taiwanese horseshoe crab.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF264067; AAF74773.1; -;  
SQ SEQUENCE 256 AA; 28517 MW; 5EC0272B88F44FF8 CRC64;

Query Match 53.2%; Score 41; DB 5; Length 256;  
Best Local Similarity 72.7%; Pred. No. 15;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 KVSASGYN 14  
| | | | |  
Db 109 KVSASGYI 119

RESULT 11  
Q9FWU1 PRELIMINARY; PRT; 377 AA.  
AC Q9FWU1;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE HYPOTHETICAL 39.3 KDA PROTEIN.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;  
OC Oryza.  
OX NCBI\_TaxID=4530;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,  
RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,  
RA Bowman C.L., Utterback T.R., Khalak H., Feldblyum T.V.,  
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;  
RT "Oryza sativa chromosome 10 BAC OSJNBa0051D19 genomic sequence.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC023240; AAG13534.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 377 AA; 39255 MW; 28E1C75270877C4D CRC64;

Query Match 53.2%; Score 41; DB 10; Length 377;  
Best Local Similarity 57.1%; Pred. No. 23;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ASKSVSASGYN 15  
| | | | |  
Db 227 AAASCHASGYVLH 240

RESULT 12  
Q91908 PRELIMINARY; PRT; 2693 AA.  
ID Q91908;  
AC Q91908;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE INOSITOL 1,4,5-TRIPHOSPHATE RECEPTOR.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
[1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=OVARY;  
RX MEDLINE=93258819; PubMed=8387895;  
RA Kume S., Muto A., Aruga J., Nakagawa T., Michikawa T., Furuichi T.,  
RA Nakade S., Okano H., Mikoshiba K.;  
RT "The Xenopus IP3 receptor: structure, function, and localization in  
RT oocytes and eggs.";  
RL Cell 73:555-570(1993).  
DR EMBL: D14400; BAA03304.1; -;  
DR InterPro; IPR000493; -;  
DR InterPro; IPR000699; -;  
DR InterPro; IPR001682; -;  
DR InterPro; IPR003608; -;  
DR Pfam; PF01365; RYDR\_ITPR.1.  
DR PRINTS; PR00779; INSP3RECEPTR.  
DR SMART; SM00472; MIR; 1.  
KW Calcium channel.  
SQ SEQUENCE 2693 AA; 306670 MW; CBF6AFEF16A753C CRC64;

Query Match 53.2%; Score 41; DB 13; Length 2693;  
Best Local Similarity 61.5%; Pred. No. 2e+02; 4; Indels 0; Gaps 0;

QY 1 RASKSVSASGYN 13  
| | | | |  
Db 1139 KAKKSESTSSYN 1151

RESULT 13  
Q9U412 PRELIMINARY; PRT; 3469 AA.  
ID Q9U412  
AC Q9U412;

DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE SANT DOMAIN PROTEIN SMTER.  
 GN SMR OR SMTER OR CG4013.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RX SEQUENCE FROM N.A.  
 RP MEDLINE=99417957; PubMed=10488333;  
 RA Tsai C.-C., Kao H.-Y., Yao T.-P., McKeown M., Evans R.M.;  
 RT "SMTER, a Drosophila nuclear receptor coregulator, reveals that Ecr-  
 mediated repression is critical for development.";  
 RL MOL. Cell 4:175-186(1999).  
 DR EMBL; AF175223; AAD52614.1; -.  
 DR FlyBase; FBgn0024308; Smr.  
 DR InterPro; IPR000104; -.  
 DR InterPro; IPR001005; -.  
 DR InterPro; IPR002086; -.  
 DR Pfam; PF00249; myb\_DNA-binding; 1.  
 DR PRINTS; PR00308; ANTIFREEZE1.  
 DR PROSITE; PS00687; ALDEHYDE\_DEHYDR\_GLU; UNKNOWN\_1.  
 DR SMART; SM00395; SANT; 1.  
 DR SEQUENCE 3469 AA; 364115 MW; 6284E14C5C247CD9 CRC64;  
 SQ  
 Query Match 53.2%; Score 41; DB 5; Length 3469;  
 Best Local Similarity 63.6%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ASKVSASGYN 12  
 Db 1208 ASKAINAEGYN 1218  
 RESULT 14  
 Q9VYJ9 ID Q9VYJ9 PRELIMINARY; PRT; 3502 AA.  
 AC Q9VYJ9  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE SMR PROTEIN.  
 GN SMR OR CG4013.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RX SEQUENCE FROM N.A.  
 RP STRAIN=BERKELEY;  
 RC MEDLINE=20196006; PubMed=10731133;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.F., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei H., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; -.  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003490; AAF48196.1; -.  
 DR FlyBase; FBgn0024308; Smr.  
 DR InterPro; IPR000104; -.  
 DR InterPro; IPR001005; -.  
 DR InterPro; IPR002086; -.  
 DR Pfam; PF00249; myb\_DNA-binding; 1.  
 DR PRINTS; PR00308; ANTIFREEZE1.  
 DR PROSITE; PS00687; ALDEHYDE\_DEHYDR\_GLU; 1.  
 DR SMART; SM00395; SANT; 1.  
 DR SEQUENCE 3502 AA; 369068 MW; 74C8004F9DA8F8D9 CRC64;  
 SQ  
 Query Match 53.2%; Score 41; DB 5; Length 3502;  
 Best Local Similarity 63.6%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ASKVSASGYN 12  
 Db 1347 ASKAINAEGYN 1357  
 RESULT 15  
 Q02240 ID Q02240 PRELIMINARY; PRT; 328 AA.  
 AC Q02240  
 DT 01-JUL-1997 (TReMBLrel. 04, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE MEC-3 PROTEIN.  
 GN MEC-3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 ON NCBI\_TaxID=6239;  
 RX SEQUENCE FROM N.A.  
 RP Wild A.;  
 RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP MEDLINE=99006961; PubMed=9851916;  
 RX none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 CC -1- SIMILARITY: TO OTHER HOMEBOX DOMAINS.  
 DR EMBL; Z81054; C2B02885.2; -.  
 DR HSSP; P14859; lFOG.  
 DR InterPro; IPR000047; -.

DR InterPro; IPR001356; -.  
 DR InterPro; IPR001781; -.  
 DR Pfam; PF00046; homeobox; 2.  
 DR Pfam; PF00412; LIM; 4.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR PRODOM; PD000094; -; 2.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS0071; HOMEBOX\_2; 1.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; UNKNOWN\_1.  
 DR PROSITE; PS00023; LIM\_DOMAIN\_2; 2.  
 DR SMART; SM00389; HOX; 1.  
 DR SMART; SM00132; LIM; 1.  
 KW DNA-binding; Homeobox; LIM motif; Metal-binding; Nuclear protein;  
 KW Zinc.  
 SQ SEQUENCE 328 AA; 37953 MW; 3287768804CDC253 CRC64;

Query Match 51.9%; Score 40; DB 5; Length 328;  
 Best Local Similarity 46.7%; Pred. No. 30;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 RASKSVSASGYNYMH 15  
 : | | | | | | | | | |  
 Db 194 QVKKEYDAYGYNFEH 208

Search completed: June 28, 2001, 16:08:27  
 Job time: 955 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:14:38 ; Search time 362.28 Seconds  
(without alignments)  
1.171 Million cell updates/sec

Title: US-09-724-406-30  
Perfect score: 31  
Sequence: 1 LASNLES 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_0601.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	31	100.0	7	18 AAW27345	CDR2 from murine a
2	31	100.0	7	21 AAB09999	H. pylori HSP60-b1
3	31	100.0	17	16 AAR85179	GM-CSF receptor al
4	31	100.0	103	15 AAR47933	Light chain region
5	31	100.0	103	20 AAW89349	Mouse 23F2G light
6	31	100.0	107	18 AAW27353	Light chain variab
7	31	100.0	108	20 AAY13949	VL chain of anti-C
8	31	100.0	110	15 AAR60810	Light chain variab
9	31	100.0	111	15 AAR53930	Light chain variab
10	31	100.0	111	15 AAR60528	Light chain variab
11	31	100.0	111	17 AAR85242	Humanised antibody

12	31	100.0	111	17 AAR85236	Murine 206 antibod
13	31	100.0	111	20 AAW89351	Humanised 23F2G 11
14	31	100.0	111	21 AAB10019	H. pylori HSP60-b1
15	31	100.0	112	12 AAR10539	Chimeric MAB 9.2.2
16	31	100.0	112	16 AAR79158	Human IgE receptor
17	31	100.0	112	16 AAR80272	VI region from an
18	31	100.0	112	16 AAR79156	Human IgE receptor
19	31	100.0	112	18 AAW27358	Light chain variab
20	31	100.0	112	18 AAW27527	Light chain variab
21	31	100.0	131	13 AAR28668	pl2-k2. Synthetic
22	31	100.0	237	20 AAW73873	Human antiFc epsil
23	31	100.0	240	20 AAW73875	Human antiFc epsil
24	31	100.0	260	16 AAR77617	Anti-C5 MAB N19/8
25	31	100.0	3096	22 AAB46770	FANCIP3 protein fr
26	29	93.5	215	15 AAR74781	Light chain of 59.
27	29	93.5	230	11 AAR06610	Intact natural cut
28	29	93.5	230	15 AAR46295	F. solani pisi pre
29	29	93.5	230	15 AAR59796	Fusarium solani pi
30	29	93.5	230	15 AAR59801	Fusarium solani pi
31	28	90.3	87	16 AAR80081	Mouse derived ligh
32	28	90.3	91	20 AAW95478	Mouse derived RT3
33	28	90.3	110	15 AAR60564	Anti-carcinoembryo
34	28	90.3	110	20 AAY39532	Murine COL1 VK cha
35	28	90.3	110	20 AAY39534	Humanised Murine C
36	28	90.3	110	20 AAY39535	Humanised Murine C
37	28	90.3	110	20 AAY39544	Humanised Murine C
38	28	90.3	110	20 AAY39546	Humanised Murine C
39	28	90.3	111	15 AAR48622	Sequence of the hu
40	28	90.3	111	15 AAR48623	Sequence of the hu
41	28	90.3	111	15 AAR65172	Murine NM-01 varia
42	28	90.3	111	16 AAR65174	Region producing h
43	28	90.3	121	14 AAR33346	Sequence of the va
44	28	90.3	121	15 AAR48615	Sequence of the mo
45	28	90.3	121	17 AAW09259	Monoclonal antibody

## ALIGNMENTS

RESULT 1  
AAW27345  
ID AAW27345 standard; peptide; 7 AA.  
XX  
AC AAW27345;  
XX  
DT 12-DEC-1997 (first entry)  
XX  
DE CDR2 from murine anti-human IgE receptor antibody heavy chain.  
XX  
KW Complementarity determining region; CDR2; murine; mouse; human;  
KW high affinity; immunoglobulin E; receptor; monoclonal antibody;  
KW IgE; MAB; heavy chain; variable region; humanised; semi-chimeric;  
KW chimeric; treatment; prevention; disease; allergy.  
XX  
OS Mus spp.  
XX  
PN JP09191886-A.  
XX  
PD 29-JUL-1997.  
XX  
PF 19-JAN-1996; 96JP-0024816.  
XX  
PR 19-JAN-1996; 96JP-0024816.  
XX  
PA (ASAK ) ASAKI BREWERIES LTD.  
PA (NIKK-) NIKKA WHISKY KK.  
PA (TORI ) TORII YAKUHN KK.  
PA (TSUR/) TSURA T.  
XX  
DR WPI; 1997-429186/40.  
XX  
PT Humanised, semi-chimeric and chimeric antibodies against human high-affinity IgE receptor - useful medicinally and have low

PT antigenicity in humans

PS Claim 1; Page 12; 26pp: Japanese.

XX The present complementarity determining region 2 (CDR2), which is  
CC from a murine, anti-human high affinity immunoglobulin E (IgE)  
CC receptor, monoclonal antibody (MAB) heavy chain variable region,  
CC can be used in the preparation of humanised or semi-chimeric  
CC anti-human high affinity IgE receptor MAB. The MAB can be used to  
CC treat or prevent diseases, specifically allergies, associated with  
CC the receptor. The humanised, semi-chimeric or chimeric MAB have  
CC very low antigenicity in humans.

XX Sequence 7 AA;

Query Match 100.0%; Score 31; DB 18; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
| | | | |  
Db 1 lasnles 7

RESULT 2

AAB09999  
ID AAB09999 standard; Protein; 7 AA.

XX AC AAB09999;

XX 01-NOV-2000 (first entry)

DE H. pylori HSP60-binding antibody light chain CDR2 protein fragment.

XX Acid-resistant microorganism; detection; faecal; intestine; infection;  
KW monoclonal antibody; heat shock protein; HSP60; light chain; CDR;  
KW complementarity determining region.

XX Unidentified.

XX WO200026671-A1.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-EP08212.

XX 29-OCT-1998; 98EP-0120517.

PR 06-NOV-1998; 98EP-0120687.

XX (CONN-) CONNEX GMBH.

XX Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;  
PI Ringels A;

XX WPI; 2000-365747/31.

DR N-PSDB; AAA40155.

XX Detecting infection by acid-fast microbes for diagnosis of Helicobacter  
PT pylori, comprises reacting a faecal sample with two binding reagents for  
PT antigens that survive intestinal passage

XX Claim 20; Page 20; 84pp: German.

XX This invention describes a novel method for the detection of a mammalian  
CC infection by an acid-resistant microorganism (A) by treating a faecal  
CC sample with at least two different monoclonal antibodies (MAB) (or their  
CC fragments or derivatives) or aptamers (collectively (I)) and detecting  
CC formation of a complex (C) between (I) and the corresponding antigen of  
CC (A). The first and second (I) bind to epitopes of different antigens  
CC (Ag). These epitopes are present, after passage through the intestines,  
CC in at least some mammals, and have either: (i) their native structure;  
CC or (ii) a structure against which an antibody is produced by an animal

CC infected or immunized with (A), or its extract, lysate, derived protein  
CC or fragment, or with a synthetic peptide. Practically all mammals display  
CC at least one of the specified epitopes. The method is used to detect  
CC infection by acid-fast bacteria, particularly of the genera Helicobacter,  
CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,  
CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used  
CC therapeutically. The method is direct and non-invasive, and provides an  
CC inexpensive and easily standardizable diagnosis, despite possible  
CC degradation of antigens during passage through the intestines. This  
CC sequence represents a fragment of a H. pylori heat shock protein,  
CC HSP60-binding antibody light chain complementarity determining region  
CC CDR1 which is used to illustrate the method of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 31; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
| | | | |  
Db 1 lasnles 7

RESULT 3

AAR85179  
ID AAR85179 standard; peptide; 17 AA.

XX AC AAR85179;

XX 06-JUN-1996 (first entry)

DE GM-CSF receptor alpha-chain binding peptide.

XX Alpha-chain; GM-CSF; receptor; granulocyte; macrophage;  
KW colony stimulating factor; binding assay; antagonist;  
KW inflammation; autoimmune disease; agonist; upregulator;  
KW immune system cell; CDR II; light chain; antibody 23.2.

XX Synthetic.

XX WO9529690-A1.

XX 09-NOV-1995.

XX 26-APR-1995; 95WO-US05160.

XX 29-APR-1994; 94US-0235404.

XX (UYPE-) UNIV PENNSYLVANIA.

XX (WIST-) WISTAR INST.

XX Kieber-Emmons T, Von Feldt JM, Weiner DB, Williams WV;

XX WPI; 1995-392919/50.

XX Identifying biologically active protein peptide mimics - by prepn.  
PT of recombinant antibody libraries and synthesising complementary  
PT determining region sequences.

XX Example 2; Page 49; 70pp: English.

XX The present peptide binds the alpha-chain of the GM-CSF receptor,  
CC and can therefore be used in binding assays in place of GM-CSF. It  
CC can also be used as an antagonist, i.e. in the treatment of  
CC inflammation or autoimmune diseases, or in the prodn. of GM-CSF  
CC agonists, i.e. upregulators for the prodn. of immune system cells.  
CC The peptide is derived from CDR II of the recombinant antibody  
CC light chain analogue of GM-CSF, 23.2.

XX Sequence 17 AA;

Query Match 100.0%; Score 31; DB 16; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 Db 2 lasnles 8

## RESULT 4

AAW47933  
 ID AAR47933 standard; Protein; 103 AA.  
 AC AAR47933;  
 XX  
 DT 24-JUL-1994 (first entry)  
 XX  
 DE Light chain region of 23F2G.  
 XX  
 KW Amplification; 23F2G; humanised antibody; heavy chain; light chain;  
 KW hybridoma; inflammation; CD18; human leukocyte integrins; mAb 60.3;  
 KW monoclonal antibody; LFA-1; adhesion; migration; multiple  
 KW sclerosis; MS.  
 XX  
 PN W09402175-A.  
 XX  
 PD 03-FEB-1994.  
 XX  
 PF 16-JUL-1993; 93WO-US06734.  
 XX  
 PR 16-JUL-1992; 92US-0915068.  
 PR 10-MAY-1993; 93US-0060699.  
 XX  
 PA (ICOS-) ICOS CORP.  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Rose LM;  
 XX  
 DR WPI: 1994-048551/06.  
 DR N-PSDB; AAQ35915.  
 XX  
 PS Antibodies immunologically reactive with the CD18 of human  
 PT leukocyte integrins and/or competing with mAb 60.3 for binding to  
 PT human LFA-1 - for alleviating symptoms associated with  
 PT inflammatory disease states  
 XX  
 PS Example 6; Page 43; 58pp; English.

CC Total RNA was isolated from the hybridoma cell line 23F2G and first  
 CC strand cDNA was synthesised using the total RNA as a template. The  
 CC first strand cDNA was used as a template for PCR to obtain double  
 CC stranded DNA fragments encoding the variable regions of both the  
 CC heavy and light chains of monoclonal antibody 23F2G. The sequence  
 CC shown is that of the light chain variable region of MAb 23F2G  
 CC The humanised form of MAb 23F2G may be administered to  
 CC alleviate symptoms associated with inflammatory disease states, esp.  
 CC for the inhibition of inflammatory processes associated with  
 CC multiple sclerosis. The MAb blocks leukocyte adhesion and  
 CC migration to inflammatory sites. The MAb is an anti-CD18 integrin  
 CC antibody which competes with MAb 60.3 for binding to LFA-1.  
 CC See also AAR47931-6.  
 XX  
 SQ Sequence 103 AA;

Query Match 100.0%; Score 31; DB 15; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 Db 46 lasnles 52

## RESULT 5

AAW89349  
 ID AAW89349 standard; Protein; 103 AA.  
 XX  
 AC AAW89349;  
 XX  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE Mouse 23F2G light chain variable region.  
 XX  
 KW Mouse; humanised; antibody; heavy chain variable region; light chain;  
 KW 23F2G; inflammatory disease; multiple sclerosis; common beta chain;  
 KW CD18; human leukocyte integrin.  
 XX  
 OS Mus sp.  
 XX  
 PN US5854070-A.  
 XX  
 PD 29-DEC-1998.  
 XX  
 PF 21-JAN-1997; 97US-0785571.  
 XX  
 PR 16-JUL-1993; 93US-0094535.  
 PR 16-JUL-1992; 92US-0915068.  
 PR 10-MAY-1993; 93US-0060699.  
 PR 28-FEB-1995; 95US-0396089.  
 PR 21-JAN-1997; 97US-0785571.  
 XX  
 PA (ICOS-) ICOS CORP.  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Rose LM;  
 XX  
 DR WPI: 1999-094920/08.  
 DR N-PSDB; AAV81907.  
 XX  
 PT Hybridomas that secrete anti-CD18 antibody - useful for treating  
 PT inflammation specifically multiple sclerosis  
 XX  
 PS Example 6; Column 25-26; 24pp; English.

CC The present invention describes the hybridoma cell line 23F2G  
 CC (ATCC HB 11801) and the Chinese hamster ovary (CHO) cell lines 6E6  
 CC (ATCC CRL 11398) and B13-24 (ATCC CRL 11397). Monoclonal antibody (MAB)  
 CC 23F2G (and its humanised variants 6E6 and B13-24) is an anti-human CD18  
 CC antibody for the treatment of inflammation specifically multiple  
 CC sclerosis. The present sequence is mouse 23F2G light chain variable  
 CC region from the present invention.  
 XX  
 SQ Sequence 103 AA;

Query Match 100.0%; Score 31; DB 20; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 Db 46 lasnles 52

## RESULT 6

AAW27353  
 ID AAW27353 standard; Protein; 107 AA.  
 XX  
 AC AAW27353;  
 XX  
 DT 12-DEC-1997 (first entry)  
 XX  
 DE Light chain variable region of human CRA2 antibody.

KW Complementarity determining region; CDR; murine; mouse; human;  
 KW high affinity; immunoglobulin E; receptor; monoclonal antibody;  
 KW IgE; MAb; light chain; variable region; humanised; semi-chimeric;  
 KW chimeric; treatment; prevention; disease; allergy; CRA2.

XX Homo sapiens.

XX JP09191886-A.

XX 29-JUL-1997.

XX 19-JAN-1996; 96JP-0024816.

XX 19-JAN-1996; 96JP-0024816.

XX (ASAK ) ASahi BREWERIES LTD.

XX (NIKK-) NIKKA WHISKEY KK.

XX (TORI ) TORII YAKUHIIN KK.

XX (TSUR/) TSURA T.

XX WPI; 1997-429186/40.

XX N-PSDB; AAT90024.

XX Humanised, semi-chimeric and chimeric antibodies against human  
 PT high-affinity IgE receptor - useful medicinally and have low  
 PT antigenicity in humans

XX Disclosure; Fig 1; 26pp; Japanese.

XX The CDNA encoding the present sequence, the light chain variable  
 CC region of the human antibody (Ab) CRA2, was used in the preparation  
 CC of a humanised or semi-chimeric monoclonal Ab (MAB), comprising  
 CC complementarity determining regions (CDR) from a murine, anti-human  
 CC high affinity immunoglobulin E (IgE) receptor, MAB. The humanised,  
 CC semi-chimeric or chimeric MAB can be used to treat or prevent  
 CC diseases, specifically allergies, associated with the receptor, and  
 CC has very low antigenicity in humans.

XX Sequence 107 AA;

Query Match 100.0%; Score 31; DB 18; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 8.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LASNLES 7  
 Db | | | | |  
 49 lasnles 55

RESULT 7

AA13949

ID AAY13949 standard; Protein; 108 AA.

XX AC AAY13949;

XX 14-JUL-1999 (first entry)

XX VL chain of anti-CD40 monoclonal antibody.

XX Inducer; antibody production; antigen-specific antibody; vaccine; HIV;  
 KW B-cell activating signal; infectious disease; therapy; diagnosis.

XX Synthetic.

XX WO9922008-A1.

XX 06-MAY-1999.

XX 09-OCT-1998; 98WO-JP04563.

XX 27-OCT-1997; 97JP-0294393.

XX

PA (SUME ) SUMITOMO ELECTRIC IND CO.

XX Sakaguchi N;

XX WPI; 1999-303020/25.

XX Inducer for production of antigen-specific antibody of high avidity

XX Example; Page 29-30; 36pp; Japanese.

XX This sequence is the VL chain of an anti-CD40 monoclonal antibody used in  
 CC the construction of an inducer of the invention.

CC The inducer is for the production of an antigen-specific antibody with  
 CC high avidity, consists of a chimeric molecule containing the antigen  
 CC together with a B-cell activating signal. The inducer is a chimeric  
 CC molecule containing: (1) an antigen portion recognised by a B-cell  
 CC antigen receptor; and (2) a signal transmitter portion, such as anti-CD40  
 CC antibody or its active fragments, which activates B-cells to produce the  
 CC antigen-specific antibody. The inducer may be incorporated into a vaccine  
 CC for prevention and treatment of infectious diseases including HIV, and  
 CC used to produce antigen-specific monoclonal antibodies for the treatment,  
 CC diagnosis and investigation of diseases.

XX Sequence 108 AA;

Query Match 100.0%; Score 31; DB 20; Length 108;

Best Local Similarity 100.0%; Pred. No. 8.2; 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0;

OY 1 LASNLES 7

Db | | | | |

50 lasnles 56

RESULT 8

AAR60810

ID AAR60810 standard; Protein; 110 AA.

XX AC AAR60810;

XX 06-DEC-1994 (first entry)

XX Light chain variable region of antibody HCMV16.

XX Antibody; humanised antibody; variable region; immune response;  
 KW cytomegalovirus; glycoprotein; complementarity determining region;  
 KW CDR; heavy chain; light chain; ds.

XX Mus musculus.

XX WO9409136-A.

XX 28-APR-1994.

XX 15-OCT-1993; 93WO-GB02134.

XX 15-OCT-1992; 92GB-0021654.

XX (KETT-) KETTACK LODGE CAMPUS 2.

XX Carr FU, Hamilton AA, Harris WJ;

XX WPI; 1994-151321/18.

XX N-PSDB; AAQ73749.

XX New humanised monoclonal antibodies against Human Cytomegalovirus  
 PT - contain murine CDR and human framework modified to increase  
 PT specificity or affinity, for treatment or prevention of HCMV  
 PT infection.

XX Disclosure; Figure 2; 53pp; English.

XX



CC Human monoclonal antibodies may be made specific for a pathogen by  
 CC substituting the complementarity determining regions (CDR's) from  
 CC the murine antibody variable domains which show desirable binding  
 CC properties to that pathogen, into the human immunoglobulin heavy and  
 CC light chain variable domains. These "humanised" antibodies should  
 CC elicit a considerably reduced immune response in humans compared to  
 CC chimeric antibodies as they contain considerably less murine  
 CC components and their half life in the circulation should approach  
 CC that of natural human antibodies. This sequence is taken from a  
 CC donor monoclonal antibody HCW16, specific for the human  
 CC cytomegalovirus gH glycoprotein from which CDR's were taken to  
 CC construct humanised antibodies. Humanised antibodies are described  
 CC in AAQ63902-Q63908.  
 XX  
 SQ Sequence 110 AA;

Query Match 100.0%; Score 31; DB 15; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 8.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 DB 54 lasnles 60

RESULT 9  
 AAR53930  
 ID AAR53930 standard; Protein; 111 AA.  
 XX  
 AC AAR53930;  
 XX  
 DT 05-DEC-1994 (first entry)  
 XX  
 DE Light chain variable region of humanised antibody 16HuVX.

XX Antibody; humanised antibody; variable region; immune response;  
 KW cytomegalovirus; glycoprotein; complementarity determining region;  
 KW CDR; heavy chain; light chain.

XX Homo sapiens.  
 OS Mus musculus.

XX WO9409136-A.

XX 28-APR-1994.

XX 15-OCT-1993; 93WO-GB02134.

XX 15-OCT-1992; 92GB-0021654.

XX (KETT-) KETTOK LODGE CAMPUS 2.

XX Carr FJ, Hamilton AA, Harris WJ;

XX WPI; 1994-151321/18.

XX N-PSDB; AAQ63908.

XX New humanised monoclonal antibodies against Human Cytomegalovirus  
 PT - contain murine CDR and human framework modified to increase  
 PT specificity or affinity, for treatment or prevention of HCMV  
 PT infection.

XX Example 1; Figure 5; 53pp; English.

XX Human monoclonal antibodies may be made specific for a pathogen by  
 CC substituting the complementarity determining regions (CDR's) from  
 CC the murine antibody variable domains which show desirable binding  
 CC properties to that pathogen, into the human immunoglobulin heavy and  
 CC light chain variable domains. These "humanised" antibodies should  
 CC elicit a considerably reduced immune response in humans compared to  
 CC chimeric antibodies as they contain considerably less murine  
 CC components and their half life in the circulation should approach

CC that of natural human antibodies. This sequence is a humanised heavy  
 CC chain variable region comprising CDR's from the donor monoclonal  
 CC antibody HCW16, specific for the human cytomegalovirus gH  
 CC glycoprotein. The framework region of the human heavy chain has been  
 CC modified. Humanised antibodies are described in AAQ63902-Q63908.  
 XX  
 SQ Sequence 111 AA;

Query Match 100.0%; Score 31; DB 15; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 8.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 DB 54 lasnles 60

RESULT 10  
 AAR60528  
 ID AAR60528 standard; Protein; 111 AA.

XX  
 AC AAR60528;

XX 07-APR-1995 (first entry)

XX Light chain variable region of monoclonal antibody 5E8.

XX Secretion sequence; vector; antibody; single chain antibody;  
 KW promoter; Bacillus subtilis; ss.

XX Synthetic.

XX EP614982-A.

XX 14-SEP-1994.

XX 05-MAR-1994; 94EP-0200564.

XX 10-MAR-1993; 93IT-OMI0456.

XX (ENIE ) ENIRICERCHE SPA.

XX Cuzzoni A, De FERRA F, Grandi G, Tortora O, Tosi C;

XX WPI; 1994-281209/35.

XX N-PSDB; AAQ71286.

XX Vector for secretion of single chain antibody from Bacillus  
 PT subtilis - contg. neutral protease promoter and specific leader  
 PT sequence, providing high yield of soluble antibody for diagnostic  
 PT or therapeutic use

XX Example 3; Figure 2; 27pp; English.

XX The sequence encoding the variable region of the light chain of the  
 CC monoclonal antibody 5E8 (specific for the alpha subunit of human  
 CC gonadotropin) was amplified for its use in a recombinant vector  
 CC which also comprised (1) the promoter of the gene for the neutral  
 CC protease of Bacillus subtilis BGSC 1A341, (2) the secretion sequence  
 CC described in AAQ71278. The amplified fragment was used to produce a  
 CC DNA sequence encoding an antibody in single molecule form with the  
 CC sequence VH/VK-L-VK/VH-(TAG)n, where VH and VK are the variable  
 CC regions of the heavy and light chains of the antibody and L is the  
 CC linker of sequence Val-Ser-Ser-(Gly(4)-Ser)<sub>3</sub>. TAG is a peptide  
 CC recognised by polyclonal antibodies directed towards the same peptide,  
 CC n is 1 or 0. The vector is used for the production of single chain  
 CC antibodies which have improved pharmacokinetic properties and can be  
 CC produced more economically than monoclonal antibodies. The vector  
 CC allows production of antibodies in completely soluble form with  
 CC secretion in high yield.

XX Sequence 111 AA;



## RESULT 13

AAW89351  
ID AAW89351 standard; Protein; 111 AA.

XX AC AAW89351;  
XX XX

XX XX 02-MAR-1999 (first entry)  
XX XX

XX XX Humanised 23F2G light chain variable region.  
XX XX

XX KW Mouse; humanised; antibody; heavy chain variable region; light chain;  
XX KW 23F2G; inflammatory disease; multiple sclerosis; common beta chain;  
XX KW CD18; human leukocyte integrin.  
XX XX

XX OS Mus sp.  
XX OS Homo sapiens.

XX OS Synthetic.  
XX XX

XX PN US5854070-A.  
XX XX

XX PD 29-DEC-1998.  
XX XX

XX PF 21-JAN-1997; 97US-0785571.  
XX XX

XX PR 16-JUL-1993; 93US-0094535.  
XX PR

XX PR 16-JUL-1992; 92US-0915068.  
XX PR

XX PR 10-MAY-1993; 93US-0060699.  
XX PR

XX PR 28-FEB-1995; 95US-0396089.  
XX PR

XX PR 21-JAN-1997; 97US-0785571.  
XX XX

XX PA (ICOS-) ICOS CORP.  
XX PA (UNIW ) UNIV WASHINGTON.

XX XX  
XX PI Rose LM;  
XX XX

XX DR WPI; 1999-094920/08.  
XX DR N-PSDB; AAV81909.

XX XX  
XX XX Hydrindomas that secrete anti-CD18 antibody - useful for treating  
XX PT Inflammation specifically multiple sclerosis  
XX XX

XX PS Example 6; Column 27-30; 24pp; English.  
XX XX

XX CC The present invention describes the hybridoma cell line 23F2G  
(ATCC HB 11801) and the Chinese hamster ovary (CHO) cell lines 6E6  
(ATCC CRL 11398) and B13-24 (ATCC CRL 11397). Monoclonal antibody (MAB)  
23F2G (and its humanised variants 6E6 and B13-24) is an anti-human CD18  
antibody for the treatment of inflammation, specifically multiple  
sclerosis. The present sequence represents humanised 23F2G light chain  
variable region from the present invention.  
XX XX

XX SQ Sequence 111 AA;

Query Match 100.0%; Score 31; DB 20; Length 111;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 54 lasnles 60

## RESULT 14

AAAB10019  
ID AAAB10019 standard; Protein; 111 AA.

XX AC AAB10019;  
XX XX

XX DT 01-NOV-2000 (first entry)  
XX XX

DE XX H. pylori HSP60-binding antibody light chain protein.  
KW Acid-resistant microorganism; detection; faecal; intestine; infection;  
KW monoclonal antibody; heat shock protein; HSP60; light chain.  
XX OS Unidentified.  
XX PN WO200026671-A1.  
XX PD 11-MAY-2000.  
XX XX 29-OCT-1999; 99WO-EP08212.  
XX PF 29-OCT-1998; 98EP-0120517.  
XX PR 06-NOV-1998; 98EP-0120687.  
XX XX (CONN-) CONNEX GMBH.  
XX PA Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;  
XX PI Ringels A;  
XX XX  
XX DR WPI; 2000-365747/31.  
XX DR N-PSDB; AAA40199.  
XX XX  
XX PT Detecting infection by acid-fast microbes for diagnosis of Helicobacter  
XX PT pylori, comprises reacting a faecal sample with two binding reagents for  
XX XX antigens that survive intestinal passage -  
XX XX Disclosure; Fig 1; 84pp; German.  
XX CC This invention describes a novel method for the detection of a mammalian  
XX CC infection by an acid-resistant microorganism (A) by treating a faecal  
XX CC sample with at least two different monoclonal antibodies (MAB) (or their  
XX CC fragments or derivatives) or aptamers (collectively (I)) and detecting  
XX CC formation of a complex (C) between (I) and the corresponding antigen of  
XX CC (A). The first and second (I) bind to epitopes of different antigens  
XX CC (Ag). These epitopes are present, after passage through the intestines,  
XX CC in at least some mammals, and have either: (i) their native structure;  
XX CC or (ii) a structure against which an antibody is produced by an animal  
XX CC infected or immunized with (A), or its extract, lysate, derived protein  
XX CC or fragment, or with a synthetic peptide. Practically all mammals display  
XX CC at least one of the specified epitopes. The method is used to detect  
XX CC infection by acid-fast bacteria, particularly of the genera Helicobacter,  
XX CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,  
XX CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used  
XX CC therapeutically. The method is direct and non-invasive, and provides an  
XX CC inexpensive and easily standardizable diagnosis, despite possible  
XX CC degradation of antigens during passage through the intestines. This  
XX CC sequence represents the H. pylori heat shock protein, HSP60-binding  
XX CC antibody (DMS ACC2356) light chain which is used to illustrate the method  
XX CC of the invention.  
XX SQ Sequence 111 AA;

Query Match 100.0%; Score 31; DB 21; Length 111;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 54 lasnles 60

RESULT 15  
AAAR10539  
ID AAAR10539 standard; Protein; 112 AA.  
XX AC AAR10539;  
XX XX  
XX DT 15-APR-1991 (first entry)  
XX XX  
XX DE Chimeric MAB 9.2.27 light chain variable region.

XX Chimeric antibodies; human glycoprotein antigen; melanoma; cancer;  
 KW Mus musculus.  
 OS  
 XX  
 XX EP411893-A.  
 PN  
 XX  
 XX 06-FEB-1991.  
 PD  
 XX 31-JUL-1990; 90EP-0308402.  
 PF  
 XX 31-JUL-1989; 89US-0387665.  
 PR  
 XX (ELIL ) ELI LILLY & CO.  
 PA  
 XX Beavers LS, Bumol TF, Gadski RA;  
 PI  
 XX WPI; 1991-038771/06.  
 DR  
 DR N-PSDB; AAQ10379.  
 XX Monoclonal antibody contg. recombinant DNA - binds to human  
 PT chondroitin sulphate proteoglycan on melanoma cells for melanoma  
 PT treatment and diagnosis  
 PT  
 XX Claim 1; page 17; 33pp; English.  
 PS  
 XX This sequence is encoded by recombinant DNA contained in vector  
 CC pTZK910. This is ligated to a DNA sequence encoding a human light  
 CC (L) chain constant (C) region, in the construction of pG9.2.27K.  
 CC This vector is used to transform host cells, in conjunction with  
 CC vector pN9.2.27G1 contg. murine heavy (H) chain variable (V) region  
 CC and human heavy chain C region. The resulting host cell expresses  
 CC the chimeric antibody 9.2.27 which is directed against proteoglycans  
 CC of human melanoma cells. This chimeric monoclonal antibody (MAb) is  
 CC useful in the diagnosis and treatment of melanoma. The use of human  
 CC C-regions avoids the problems associated with murine monoclonals e.g.  
 CC rapid clearance from the bloodstream due to anti-self recognition.  
 CC See also AAQ10380-84.  
 XX  
 XX Sequence 112 AA;  
 SQ

Query Match 100.0%; Score 31; DB 12; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 8.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LASNLES 7  
 Db 54 lasnles 60

Search completed: June 28, 2001, 16:14:39  
 Job time: 1327 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:01:15 ; Search time 138.34 Seconds  
(without alignments)  
1.019 Million cell updates/sec

Title: US-09-724-406-30  
Perfect score: 31  
Sequence: 1 LASNLES 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/1aa/PTCUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	17	US-08-712-212-5	Sequence 5, Appli
2	31	100.0	17	PCT-US95-05160-5	Sequence 5, Appli
3	31	100.0	111	US-08-207-169A-4	Sequence 4, Appli
4	31	100.0	130	US-08-137-117D-86	Sequence 86, Appli
5	31	100.0	130	US-08-436-717-86	Sequence 86, Appli
6	31	100.0	131	US-08-137-117D-25	Sequence 25, Appli
7	31	100.0	131	US-08-436-717-25	Sequence 25, Appli
8	28	90.3	91	US-08-273-146-49	Sequence 49, Appli
9	28	90.3	110	US-08-017-570-2	Sequence 2, Appli
10	28	90.3	110	US-08-471-426-2	Sequence 2, Appli
11	28	90.3	110	PCT-US94-01709-2	Sequence 2, Appli
12	28	90.3	111	US-08-111-080-30	Sequence 30, Appli
13	28	90.3	111	US-08-111-080-32	Sequence 32, Appli
14	28	90.3	111	US-08-273-053-11	Sequence 11, Appli
15	28	90.3	111	US-08-273-053-14	Sequence 14, Appli
16	28	90.3	111	US-08-211-980-30	Sequence 30, Appli
17	28	90.3	111	US-08-211-980-32	Sequence 32, Appli
18	28	90.3	111	PCT-US93-07967-30	Sequence 30, Appli
19	28	90.3	111	PCT-US93-07967-32	Sequence 32, Appli
20	28	90.3	121	US-08-111-080-18	Sequence 18, Appli
21	28	90.3	121	US-08-111-080-22	Sequence 22, Appli
22	28	90.3	121	US-08-211-980-18	Sequence 18, Appli
23	28	90.3	121	US-08-211-980-22	Sequence 22, Appli
24	28	90.3	121	PCT-US92-07111-17	Sequence 17, Appli
25	28	90.3	121	PCT-US93-07967-18	Sequence 18, Appli
26	28	90.3	121	PCT-US93-07967-22	Sequence 22, Appli
27	28	90.3	1618	US-08-462-467B-4	Sequence 4, Appli

28 28 90.3 2887 4 US-08-462-467B-2 Sequence 2, Appli  
29 28 90.3 2887 4 US-08-462-467B-8 Sequence 8, Appli  
30 27 87.1 6 1 US-08-137-117D-137 Sequence 137, App  
31 27 87.1 6 2 US-08-436-717-137 Sequence 137, App  
32 27 87.1 7 1 US-08-318-970B-16 Sequence 16, Appl  
33 27 87.1 7 2 US-08-483-635-18 Sequence 18, Appl  
34 27 87.1 7 2 US-08-483-632-18 Sequence 18, Appl  
35 27 87.1 37 4 US-08-767-128-40 Sequence 40, Appl  
36 27 87.1 41 3 US-08-984-277-6 Sequence 6, Appli  
37 27 87.1 64 2 US-08-765-179B-10 Sequence 10, Appl  
38 27 87.1 106 1 US-08-202-047-26 Sequence 26, Appl  
39 27 87.1 106 3 US-08-964-690-26 Sequence 26, Appl  
40 27 87.1 106 3 US-08-466-151-6 Sequence 6, Appli  
41 27 87.1 110 1 US-08-442-542-8 Sequence 8, Appli  
42 27 87.1 110 3 US-08-483-749A-20 Sequence 20, Appl  
43 27 87.1 110 3 US-08-765-469-8 Sequence 8, Appli  
44 27 87.1 111 1 US-07-634-278-46 Sequence 46, Appl  
45 27 87.1 111 1 US-07-634-278-47 Sequence 47, Appl

ALIGNMENTS

RESULT 1  
US-08-712-212-5  
; Sequence 5, Application US/08712212  
; Patent No. 5837460  
; GENERAL INFORMATION:  
; APPLICANT: Williams, William V.  
; APPLICANT: Kieber-Emmons, Thomas  
; APPLICANT: Weiner, David B.  
; APPLICANT: Vonfeldt, Joan M.  
; TITLE OF INVENTION: Biologically active peptides and  
; TITLE OF INVENTION: methods of identifying the same  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5837460ris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/712,212  
; FILING DATE: 03-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/235,404  
; FILING DATE: 29-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: UPN-1554  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-712-212-5

Query Match 100.0%; Score 31; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LASNLES 7  
| | | | |  
Db 2 LASNLES 8

RESULT 2  
PCT-US95-05160-5  
; Sequence 5, Application PC/TUS9505160  
; GENERAL INFORMATION:  
; APPLICANT: Williams, William V.  
; APPLICANT: Kieber-Emmons, Thomas  
; APPLICANT: Weiner, David B.  
; APPLICANT: VonFeldt, Joan M.  
; TITLE OF INVENTION: Biologically active peptides and  
; TITLE OF INVENTION: methods of identifying the same  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place, 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05160  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/235,404  
; FILING DATE: 29-APR-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: UPN-2245  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-05160-5

Query Match 100.0%; Score 31; DB 5; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LASNLES 7  
| | | | |  
Db 2 LASNLES 8

RESULT 3  
US-08-207-169A-4  
; Sequence 4, Application US/08207169A  
; Patent No. 5674712  
; GENERAL INFORMATION:  
; APPLICANT: GRANDI, GUIDO  
; APPLICANT: DE FERRA, FRANCESCA  
; APPLICANT: TOSI, CLAUDIO  
; APPLICANT: TORTORA, ORNELLA

; APPLICANT: CUZZONI, ANNA  
; TITLE OF INVENTION: RECOMBINANT VECTOR AND USE THEREOF FOR  
; TITLE OF INVENTION: EXOCELLULAR PREPARATION OF SINGLE MOLECULA ANTIBODIES FROM  
; TITLE OF INVENTION: BACILLUS SUBTILIS  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; ADDRESS: 1755 S. Jefferson Davis Highway, Suite 400  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/207,169A  
; FILING DATE: 08-MAR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5674712man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 2264-061-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-207-169A-4

Query Match 100.0%; Score 31; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LASNLES 7  
| | | | |  
Db 54 LASNLES 60

RESULT 4  
US-08-137-117D-86  
; Sequence 86, Application US/08137117D  
; Patent No. 5795965  
; GENERAL INFORMATION:  
; APPLICANT: TSUCHIYA, Masayuki  
; APPLICANT: SATO, Koh  
; APPLICANT: BENDIG, Mary  
; APPLICANT: JONES, Steven  
; APPLICANT: SALDANHA, Jose  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
; NUMBER OF SEQUENCES: 158  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/137,117D  
;; FILING DATE: 20-DEC-1993  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO PCT/JP92/00544  
;; FILING DATE: 24-APR-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 4-32084  
;; FILING DATE: 19-FEB-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 3-95476  
;; FILING DATE: 25-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: WEGNER, Harold C.  
;; REGISTRATION NUMBER: 25,258  
;; REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
;; TELEPHONE: (202)672-5300  
;; TELEFAX: (202)672-5399  
;; TELEX: 904136  
;; INFORMATION FOR SEQ ID NO: 86:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 130 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-137-117D-86

Query Match 100.0%; Score 31; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 73 LASNLES 79

RESULT 5  
US-08-436-717-86  
; Sequence 86, Application US/08436717  
; Patent No. 5817790  
; GENERAL INFORMATION:  
; APPLICANT: TSUCHIYA, Masayuki  
; APPLICANT: SATO, Koh  
; APPLICANT: BENDIG, Mary  
; APPLICANT: JONES, Steven  
; APPLICANT: SALDANHA, Jose  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; NUMBER OF SEQUENCES: 158  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/436,717  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/137,117  
; FILING DATE: 20-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; APPLICATION NUMBER: WO PCT/JP92/00544  
; FILING DATE: 24-APR-1992

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 4-32084  
;; FILING DATE: 19-FEB-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 3-95476  
;; FILING DATE: 25-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: WEGNER, Harold C.  
;; REGISTRATION NUMBER: 25,258  
;; REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
;; TELEPHONE: (202)672-5300  
;; TELEFAX: (202)672-5399  
;; TELEX: 904136  
;; INFORMATION FOR SEQ ID NO: 86:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 130 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-436-717-86

Query Match 100.0%; Score 31; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 73 LASNLES 79

RESULT 6  
US-08-137-117D-25  
; Sequence 25, Application US/08137117D  
; Patent No. 5795965  
; GENERAL INFORMATION:  
; APPLICANT: TSUCHIYA, Masayuki  
; APPLICANT: SATO, Koh  
; APPLICANT: BENDIG, Mary  
; APPLICANT: JONES, Steven  
; APPLICANT: SALDANHA, Jose  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; NUMBER OF SEQUENCES: 158  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/137,117D  
; FILING DATE: 20-DEC-1993  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/JP92/00544  
; FILING DATE: 24-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 4-32084  
; FILING DATE: 19-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 3-95476  
; FILING DATE: 25-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEGNER, Harold C.  
; REGISTRATION NUMBER: 25,258

; REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 131 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-137-117D-25

Query Match 100.0%; Score 31; DB 1; Length 131;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASNLES 7  
Db 74 LASNLES 80

RESULT 7  
US-08-436-717-25  
; Sequence 25, Application US/08436717  
; Patent No. 5817790  
; GENERAL INFORMATION:  
; APPLICANT: TSUCHIYA, Masayuki  
; APPLICANT: SATO, Koh  
; APPLICANT: BENDIG, Mary  
; APPLICANT: JONES, Steven  
; APPLICANT: SALDANHA, Jose  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
; NUMBER OF SEQUENCES: 158  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/436,717  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/137,117  
; FILING DATE: 20-DEC-1993  
; APPLICATION NUMBER: WO PCT/JP92/00544  
; FILING DATE: 24-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 4-32084  
; FILING DATE: 19-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 3-95476  
; FILING DATE: 25-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEGNER, Harold C.  
; REGISTRATION NUMBER: 25,258  
; REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 131 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-436-717-25

Query Match 100.0%; Score 31; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASNLES 7  
Db 74 LASNLES 80

RESULT 8  
US-08-273-146-49  
; Sequence 49, Application US/08273146  
; Patent No. 5855885  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Rodger  
; APPLICANT: McCafferty, John  
; APPLICANT: Chiswell, David  
; APPLICANT: Darsley, Michael J.  
; APPLICANT: Fitzgerald, Kevin  
; APPLICANT: Kenten, John H.  
; APPLICANT: Martin, Mark T.  
; APPLICANT: Titmas, Richard C.  
; APPLICANT: Williams, Richard O.  
; TITLE OF INVENTION: The Isolation and Production of  
; TITLE OF INVENTION: Catalytic Antibodies using Phage Technology  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IGEN, Inc.  
; STREET: 1530 East Jefferson St.  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20852  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/273,146  
; FILING DATE: 14-JUL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ryan, John W.  
; REGISTRATION NUMBER: 33,771  
; REFERENCE/DOCKET NUMBER: 09000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-984-8000  
; TELEFAX: 301-230-0158  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 91 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-273-146-49

Query Match 90.3%; Score 28; DB 2; Length 91;  
Best Local Similarity 85.7%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASNLES 7  
Db 46 LASNLES 52



```

RESULT 9
US-08-017-570-2
; Sequence 2, Application US/08017570
; Patent No. 5472693
; GENERAL INFORMATION:
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: MEZES, PETER S
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Duane C. Ulmer
; STREET: P.O. Box 1967
; CITY: Midland
; STATE: MI
; COUNTRY: US
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/017,570
; FILING DATE: 19930216
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-38,777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-017-570-2

Query Match 90.3%; Score 28; DB 1; Length 110;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLS 7
Db 54 LASNLQS 60
|||||:1

RESULT 10
US-08-471-426-2
; Sequence 2, Application US/08471426
; Patent No. 5806033
; GENERAL INFORMATION:
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: MEZES, PETER S
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Duane C. Ulmer
; STREET: P.O. Box 1967
; CITY: Midland
; STATE: MI
; COUNTRY: US

```

```

; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,426
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/017,570
; FILING DATE: 16-FEB-1993
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-38,777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-471-426-2

Query Match 90.3%; Score 28; DB 1; Length 110;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLS 7
Db 54 LASNLQS 60
|||||:1

RESULT 11
PCT-US94-01709-2
; Sequence 2, Application PC/TUS9401709
; GENERAL INFORMATION:
; APPLICANT: THE DOW CHEMICAL COMPANY
; APPLICANT: U.S.A. DEPT. OF HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: A NOVEL FAMILY OF ANTI-CARCINOEMBRYONIC
; TITLE OF INVENTION: ANTIGEN CHIMERIC ANTIBODIES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Duane C. Ulmer
; STREET: P.O. Box 1967
; CITY: Midland
; STATE: MI
; COUNTRY: US
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01709
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: 38,777-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

; MOLECULE TYPE: protein  
PCT-US94-01709-2

Query Match 90.3%; Score 28; DB 5; Length 110;  
Best Local Similarity 85.7%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 54 VASNLES 60

RESULT 12  
US-08-111-080-30  
; Sequence 30, Application 08/111080  
; Patent No. 5558865  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; TITLE OF INVENTION: HIV Immunotherapeutics  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: 08/111,080  
; FILING DATE:

; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/748,562  
; FILING DATE: 22-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/07111  
; FILING DATE: 24-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/039,457  
; FILING DATE: 22-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Borun, Michael F.  
; REGISTRATION NUMBER: 25,447  
; REFERENCE/DOCKET NUMBER: 31629  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-111-080-30

Query Match 90.3%; Score 28; DB 1; Length 111;  
Best Local Similarity 85.7%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 54 VASNLES 60

RESULT 13  
US-08-111-080-32  
; Sequence 32, Application 08/111080  
; Patent No. 5558865  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; TITLE OF INVENTION: HIV Immunotherapeutics  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: 08/111,080  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/748,562  
; FILING DATE: 22-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/07111  
; FILING DATE: 24-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/039,457  
; FILING DATE: 22-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Borun, Michael F.  
; REGISTRATION NUMBER: 25,447  
; REFERENCE/DOCKET NUMBER: 31629  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-111-080-32

Query Match 90.3%; Score 28; DB 1; Length 111;  
Best Local Similarity 85.7%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 54 VASNLES 60

RESULT 14  
US-08-275-053-11  
; Sequence 11, Application US/08275053  
; Patent No. 5607847  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Recombinant human anti-human immunodeficiency  
; virus antibody.  
; NUMBER OF SEQUENCES: 16  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/275,053  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB93/01798  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-275-053-11

Query Match 90.3%; Score 28; DB 1; Length 111;  
Best Local Similarity 85.7%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

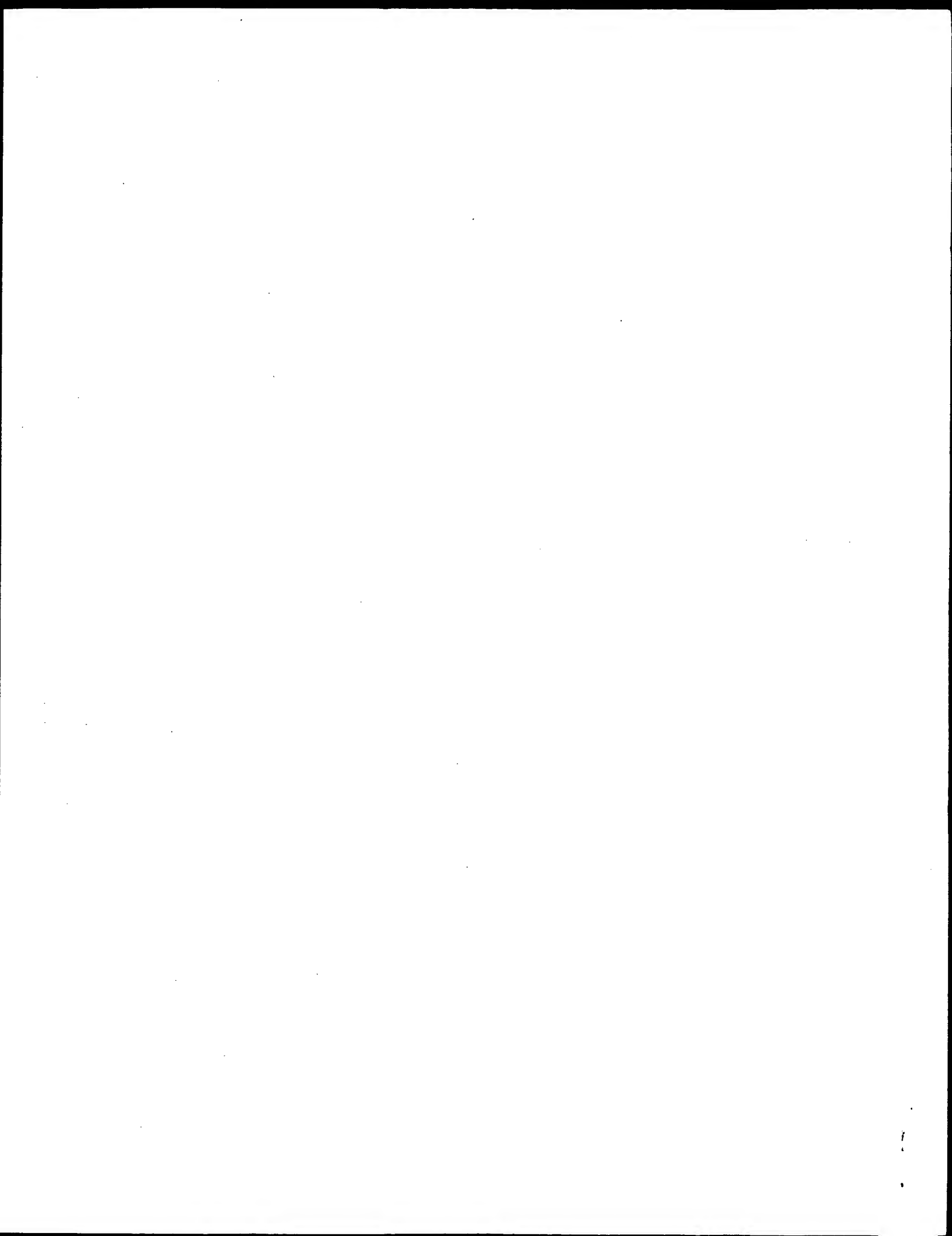
Qy 1 LASNLES 7  
Db 54 VASNLES 60

RESULT 15  
US-08-275-053-14  
; Sequence 14, Application US/08275053  
; Patent No. 5607847  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Recombinant human anti-human immunodeficiency  
; TITLE OF INVENTION: virus antibody.  
; NUMBER OF SEQUENCES: 16  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA: US/08/275,053  
; APPLICATION NUMBER: US/08/275,053  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB93/01798  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-275-053-14

Query Match 90.3%; Score 28; DB 1; Length 111;  
Best Local Similarity 85.7%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASNLES 7  
Db 54 VASNLES 60

Search completed: June 28, 2001, 16:01:15  
Job time: 523 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:58:46 ; Search time 234.85 Seconds  
(without alignments)  
2.270 Million cell updates/sec

Title: US-09-724-406-30  
Perfect score: 31  
Sequence: 1 LASNLES 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68: \*  
-: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	95	2 S25177	Ig kappa chain V r
2	31	100.0	102	2 PH1076	Ig light chain V r
3	31	100.0	111	1 KVM584	Ig kappa chain V r
4	31	100.0	111	1 KVM575	Ig kappa chain V r
5	31	100.0	111	1 KVM540	Ig kappa chain V r
6	31	100.0	111	2 D45722	anti-glycoprotein
7	31	100.0	112	2 S45715	Ig kappa chain V r
8	31	100.0	131	1 KVM5M6	Ig kappa chain pre
9	31	100.0	210	2 A56169	Ig kappa chain V r
10	31	100.0	218	2 JC5810	monoclonal antibod
11	29	93.5	94	2 PS0349	myosin heavy chain
12	29	93.5	230	1 UVFUS	cutinase (EC 3.1.-
13	29	93.5	741	2 S39082	myosin heavy chain
14	29	93.5	936	2 S39083	myosin heavy chain
15	29	93.5	955	2 S23488	myosin heavy chain
16	29	93.5	1938	1 JX0178	myosin heavy chain
17	29	93.5	1940	2 A29320	myosin heavy chain
18	28	90.3	81	2 S42193	Ig kappa chain V r
19	28	90.3	186	2 H75004	hypothetical prote
20	28	90.3	186	2 A71158	hypothetical prote
21	28	90.3	218	2 S68241	Ig kappa chain V r
22	28	90.3	367	2 T28892	hypothetical prote
23	28	90.3	376	2 T19266	hypothetical prote
24	28	90.3	1116	2 S63397	probable membrane
25	28	90.3	1225	1 B64234	hypothetical prote
26	28	90.3	1549	2 S50705	hypothetical prote
27	27	87.1	65	2 C38601	Ig kappa chain V r
28	27	87.1	65	2 B38601	Ig kappa chain V r
29	27	87.1	91	2 S25462	Ig kappa chain V r

```

30      27      87.1      93      2      A38601      Ig kappa chain V r
31      27      87.1      101      2      S59640      Ig light chain V r
32      27      87.1      102      2      PC6027      acetylcholine recep
33      27      87.1      102      2      PH1077      Ig light chain V r
34      27      87.1      107      2      S26343      Ig kappa chain V r
35      27      87.1      107      2      S26344      Ig kappa chain V r
36      27      87.1      108      1      K1HUAU      Ig kappa chain V-I
37      27      87.1      108      1      KVM554      Ig kappa chain V r
38      27      87.1      108      2      PH0092      Ig kappa chain V r
39      27      87.1      110      1      KVM510      Ig kappa chain V r
40      27      87.1      110      2      S24288      Ig kappa chain V r
41      27      87.1      111      1      KVM537      Ig kappa chain V r
42      27      87.1      111      1      KVM5C1      Ig kappa chain V r
43      27      87.1      111      1      KVM543      Ig kappa chain V r
44      27      87.1      111      1      KVM583      Ig kappa chain V r
45      27      87.1      111      1      KVM508      Ig kappa chain V r

```

## ALIGNMENTS

RESULT 1

S25177  
Ig kappa chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jan-2000  
C:Accession: S25177; S33134  
R:Monestier, M.; Fasy, T.M.; Losman, M.J.; Novick, K.E.; Muller, S.  
A:Description: Structure and binding properties of monoclonal antibodies to core hist  
A:Reference number: S25174  
A:Accession: S25177  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-95 <MON>  
A:Cross-references: EMBL:X67625; NID:g52146; PIDN:CAA47883.1; PID:g938262  
A:Accession: S33134  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-95 <MO2>  
A:Cross-references: EMBL:X67625; NID:g52146; PIDN:CAA47883.1; PID:g938262  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:16-93/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
DB 53 LASNLES 59

RESULT 2

PH1076  
Ig light chain V region (clone 74-c2) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C:Accession: PH1076  
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective  
A:Reference number: PH0971; MUID:92381444  
A:Accession: PH1076  
A:Molecule type: mRNA  
A:Residues: 1-102 <TIL>  
A:Experimental source: B cell, strain [NZB x NZW]F1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:16-94/Domain: immunoglobulin homology <IMM>

```

Query Match      100.0%; Score 31; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7
    |||||
Db 54 LASNLES 60

RESULT 3
KVMS84
Ig kappa chain V region (PC6684) - mouse (tentative sequence)
C:Species: Mus musculus (house mouse)
C:Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 31-Mar-2000
C:Accession: A01938
R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.
Nature 276, 785-790, 1978
A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.
A:Reference number: A93204; MUID:79073152
A:Accession: A01938
A:Molecule type: protein
A:Residues: 1-111 <WEI>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer
F:16-94/Domain: immunoglobulin homology <IMM>
F:23-92/Disulfide bonds: #status predicted

Query Match      100.0%; Score 31; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7
    |||||
Db 54 LASNLES 60

RESULT 4
KVMS75
Ig kappa chain V region (PC7175) - mouse (tentative sequence)
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 31-Mar-2000
C:Accession: B01938
R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.
Nature 276, 785-790, 1978
A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.
A:Reference number: A93204; MUID:79073152
A:Accession: B01938
A:Molecule type: protein
A:Residues: 1-111 <WEI>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer
F:16-94/Domain: immunoglobulin homology <IMM>
F:23-92/Disulfide bonds: #status predicted

Query Match      100.0%; Score 31; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7
    |||||
Db 54 LASNLES 60

RESULT 5
KVMS40

```

```

Ig kappa chain V region (PC7940) - mouse (tentative sequence)
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 31-Mar-2000
C:Accession: C01938; A01938
R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.
Nature 276, 785-790, 1978
A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.
A:Reference number: A93204; MUID:79073152
A:Accession: C01938
A:Molecule type: protein
A:Residues: 1-111 <WEI>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-94/Domain: immunoglobulin homology <IMM>
F:23-92/Disulfide bonds: #status predicted

Query Match      100.0%; Score 31; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7
    |||||
Db 54 LASNLES 60

RESULT 6
D45722
anti-glycoprotein H monoclonal antibody light-chain variable domain (Mab 109) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
C:Accession: D45722
R:Simpson, J.A.; Chow, J.C.; Baker, J.; Avdalovic, N.; Yuan, S.; Au, D.; Co, M.S.; Va
J. Virol. 67, 489-496, 1993
A:Title: Neutralizing monoclonal antibodies that distinguish three antigenic sites on
A:Reference number: A45722; MUID:93100833
A:Accession: D45722
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-111 <SIM>
A:Note: sequence extracted from NCBI backbone (NCBIP:120592)
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: glycoprotein
F:16-94/Domain: immunoglobulin homology <IMM>

Query Match      100.0%; Score 31; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7
    |||||
Db 54 LASNLES 60

RESULT 7
S45715
Ig kappa chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 14-Nov-1997 #text_change 07-May-1999
C:Accession: S45715
R:Kim, H.H.; Kato, K.; Yamato, S.; Igarashi, T.; Matsunaga, C.; Ohtsuka, H.; Higuchi,
FEBS Lett. 346, 246-250, 1994
A:Title: Application of (13)C NMR spectroscopy to paratope mapping for larger antigen
A:Reference number: S45714; MUID:94283606
A:Accession: S45715
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-112 <KIM>
A:Experimental source: cell line Ex-3C7
C:Superfamily: immunoglobulin V region; immunoglobulin homology

```

C:Keywords: heterotetramer; immunoglobulin  
F:16-94/Domain: immunoglobulin homology <IMM>  
F:23-92/Disulfide bonds: #status predicted

Query Match 100.0%; Score 31; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LASNLES 7  
Db 54 LASNLES 60

## RESULT 8

KVMSM6  
Ig kappa chain precursor V regions (M63, AB22, PC9245, PC4050) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 21-Jan-2000  
C:Accession: B90412; B90374; C93822; C93204; D93204; PH1078; A01935  
R:Burstein, Y.; Schechter, I.  
Biochemistry 17, 2392-2400, 1978  
A:Title: Primary structures of N-terminal extra peptide segments linked to the variable expression of immunoglobulin genes.  
A:Reference number: A90412; MUID:78235887  
A:Contents: M63  
A:Accession: B90412  
A:Molecule type: protein  
A:Residues: 1-35 <BUR>  
R:McKean, D.; Potter, M.; Hood, L.  
Biochemistry 12, 760-771, 1973  
A:Title: Mouse immunoglobulin chains. Pattern of sequence variation among kappa chains  
A:Reference number: A90374; MUID:73140225  
A:Contents: M63  
A:Accession: B90374  
A:Molecule type: protein  
A:Residues: 21-46, 'Q', 48-53, 'B', 55-57, 'Z', 59-86, 'F', 88-131 <MCK>  
A:Note: this sequence has since been revised in reference A93822  
R:McKean, D.J.; Bell, M.; Potter, M.  
Proc. Natl. Acad. Sci. U.S.A. 75, 3913-3917, 1978  
A:Title: Mechanisms of antibody diversity: multiple genes encode structurally related mo  
A:Reference number: A93822; MUID:79012520  
A:Contents: M63; AB22  
A:Accession: B93822  
A:Molecule type: protein  
A:Residues: 1-53; 69-107 <MC2>  
A:Accession: C93822  
A:Molecule type: protein  
A:Residues: 21-119, 'Y', 121-131 <MC3>  
R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
Nature 276, 785-790, 1978  
A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
A:Reference number: A93204; MUID:79073152  
A:Contents: PC9245; PC4050  
A:Accession: C93204  
A:Molecule type: protein  
A:Residues: 21-119, 'Y', 121-131 <WE1>  
A:Accession: D93204  
A:Molecule type: protein  
A:Residues: 21-119, 'L', 121-123, 'A', 125-129, 'L', 131 <WE2>  
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B c  
A:Reference number: PH0971; MUID:92381444  
A:Accession: PH1078  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 21-122 <TIL>  
A:Experimental source: B cell, strain [NZB x NZW]F1, clone 17p.73  
C:Comment: The M63 precursor sequence is shown.  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin  
F:1-20/Domain: signal sequence #status experimental <SIG>  
F:21-131/Domain: Ig kappa chain precursor V region #status experimental <MAT>  
F:36-114/Domain: immunoglobulin homology <IMM>  
F:43-112/Disulfide bonds: #status predicted

Query Match 100.0%; Score 31; DB 1; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 74 LASNLES 80

## RESULT 9

A56169  
Ig kappa chain V region (clone 23.2) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 11-Jan-2000  
C:Accession: A56169  
R:Monfardini, C.; Kleber-Emmons, T.; VonFeldt, J.M.; O'Malley, B.; Rosenbaum, H.; God  
J. Biol. Chem. 270, 6628-6638, 1995  
A:Title: Recombinant antibodies in bioactive peptide design.  
A:Reference number: A56169; MUID:95204454  
A:Accession: A56169  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-210 <MON>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 31; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 54 LASNLES 60

## RESULT 10

JC5810  
monoclonal antibody 13-1 light chain - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jan-2000  
C:Accession: JC5810  
R:Akashi, S.; Kato, K.; Torizawa, T.; Dohmae, N.; Yamaguchi, H.; Kamachi, M.; Harada,  
Biochem. Biophys. Res. Commun. 240, 566-572, 1997  
A:Title: Structural characterization of mouse monoclonal antibody 13-1 against a porp  
A:Reference number: JC5810; MUID:98063277  
A:Accession: JC5810  
A:Molecule type: protein  
A:Residues: 1-218 <AKA>  
C:Comment: This catalytic antibody has peroxidase oxidase. It is directed against a p  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:16-94/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 218;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
Db 54 LASNLES 60

## RESULT 11

PS0349  
myosin heavy chain, pectoralis profundus - chicken (fragments)

C:Species: Gallus gallus (chicken)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 13-Feb-1998  
 C:Accession: PS0349  
 R:Matsuzono, K.; Nagata, S.; Ichikawa, M.; Matsuda, G.  
 Nagasaki Igakkai Zasshi 66, 311-322, 1991  
 A:Title: Structural analysis of the chicken myosin heavy chain (comparison M. pectoralis)  
 A:Reference number: PS0349  
 A:Accession: PS0349  
 A:Molecule type: protein  
 A:Residues: 1-94 <MAT>  
 A:Experimental source: adult  
 A:Note: article in Japanese  
 A:Note: this peptide is obtained from rod portion  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: ATP

Query Match 93.5%; Score 29; DB 2; Length 94;  
 Best Local Similarity 85.7%; Pred. No. 5.6;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 DB 63 LASNIES 69

RESULT 12  
 UVFUS  
 cutinase (EC 3.1.1.-) precursor - fungus (Fusarium solani)  
 C:Species: Fusarium solani f.sp. pisi  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 11-Jun-1999  
 C:Accession: A32836; A00731; B00731; A61421  
 R:Soliday, C.L.; Dickman, M.B.; Kolattukudy, P.E.  
 J. Bacteriol. 171, 1942-1951, 1989  
 A:Title: Structure of the cutinase gene and detection of promoter activity in the 5'-fla  
 A:Reference number: A32836; MUID:89197761  
 A:Accession: A32836  
 A:Molecule type: DNA  
 A:Residues: 1-47, 'A', 49-93, 'A', 95-230 <SO1>  
 A:Cross-references: GB:M29759; NID:g168147; PIDN:AAA33335.1; PID:g168148  
 A:Note: the authors translated the codon GCA for residue 94 as Arg  
 R:Soliday, C.L.; Flurkey, W.H.; Okita, T.W.; Kolattukudy, P.E.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 3939-3943, 1984  
 A:Title: Cloning and structure determination of cDNA for cutinase, an enzyme involved in  
 A:Reference number: A00731  
 A:Accession: A00731  
 A:Molecule type: mRNA  
 A:Residues: 1-230 <SO2>  
 A:Cross-references: GB:X02640; NID:g168145; PIDN:AAA33334.1; PID:g168146  
 A:Accession: B00731  
 A:Molecule type: protein  
 A:Residues: 57-94; 113-142; 183-192 <SO3>  
 R:Soliday, C.L.; Kolattukudy, P.E.  
 Biochem. Biophys. Res. Commun. 114, 1017-1022, 1983  
 A:Title: Primary structure of the active site region of fungal cutinase, an enzyme invol  
 A:Reference number: A61421; MUID:83308716  
 A:Accession: A61421  
 A:Molecule type: protein  
 A:Residues: 113-142 <SO4>  
 R:Lin, T.S.; Kolattukudy, P.E.  
 Eur. J. Biochem. 106, 341-351, 1980  
 A:Title: Structural studies on cutinase, a glycoprotein containing novel amino acids and  
 A:Reference number: A44665; MUID:80245930  
 A:Contents: annotation; identification of glucuronylated amino end  
 C:Comment: This enzyme catalyzes the hydrolysis of cutin, a polyester that forms the str  
 C:Comment: Southern blot results suggest that the genome contains two copies of the cuti  
 C:Genetics:  
 A:Introns: 64/3  
 C:Superfamily: cutinase  
 C:Keywords: blocked amino end; glycoprotein; hydrolase  
 F:1-31/Domain: signal sequence #status predicted <SIG>  
 F:32-230/Product: cutinase #status predicted <MAT>  
 F:32/Modified site: glucuronylated amino end (Gly) (in mature form) #status experimental

F:125-187/Disulfide bonds: #status experimental  
 F:136,204/Active site: Ser, His #status experimental

Query Match 93.5%; Score 29; DB 1; Length 230;  
 Best Local Similarity 85.7%; Pred. No. 15;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 DB 71 TASNLES 77

RESULT 13  
 S39082  
 myosin heavy chain, embryonic - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 13-Aug-1999  
 C:Accession: S39082; S24349; A30170; S01265  
 R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.  
 submitted to the EMBL Data Library, August 1991  
 A:Description: Analysis of the chicken fast myosin heavy chain family: Localization o  
 A:Reference number: S39081  
 A:Accession: S39082  
 A:Molecule type: mRNA  
 A:Residues: 1-741 <MOO>  
 A:Cross-references: EMBL:M74086  
 A:Experimental source: clone Cemb3  
 R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.  
 J. Mol. Biol. 225, 1143-1151, 1992  
 A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isof  
 A:Reference number: S24348; MUID:92309413  
 A:Accession: S24349  
 A:Molecule type: mRNA  
 A:Residues: 1-12, 14-741 <MOO2>  
 A:Cross-references: EMBL:M74086  
 A:Experimental source: clone Cemb3  
 R:Ragrutta, A.A.; McCarthy, J.G.; Scherzinger, C.A.; Heywood, S.M.  
 DNA 8, 39-50, 1989  
 A:Title: Identification and developmental expression of a novel embryonic myosin heav  
 A:Reference number: A30170; MUID:89210285  
 A:Accession: A30170  
 A:Molecule type: DNA  
 A:Residues: 723-741 <LAG>  
 A:Cross-references: GB:M24691; NID:g341219; PIDN:AAA48950.1; PID:g531186  
 R:McCarthy, J.G.; Heywood, S.M.  
 Nucleic Acids Res. 15, 8069-8085, 1987  
 A:Title: A long polypurine/polypurine tract induces an altered DNA conformation o  
 A:Reference number: S01265; MUID:88040428  
 A:Accession: S01265  
 A:Molecule type: DNA  
 A:Residues: 723-741 <MCC>  
 A:Cross-references: EMBL:X06251; NID:g63600; PIDN:CAA29593.1; PID:g63601  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: actin binding; ATP; coiled coil; muscle

Query Match 93.5%; Score 29; DB 2; Length 741;  
 Best Local Similarity 85.7%; Pred. No. 55;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 DB 36 LASNMES 42

RESULT 14  
 S39083  
 myosin heavy chain, neonatal [similarity] - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 20-Jun-2000  
 C:Accession: S39083; S24350; A26821  
 R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.



submitted to the EMBL Data Library, August 1991  
 A:Description: Analysis of the chicken fast myosin heavy chain family: Localization of isoform  
 A:Reference number: S39081  
 A:Accession: S39083  
 A:Molecule type: mRNA  
 A:Residues: 1-936 <MOOI>  
 A:Cross-references: EMBL:M74087  
 R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.  
 J. Mol. Biol. 225, 1143-1151, 1992  
 A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isoform  
 A:Reference number: S24348; MUID:92309413  
 A:Accession: S24350  
 A:Molecule type: mRNA  
 A:Residues: 3-466, 'O', 468-641, 'R', 643-936 <MOO2>  
 A:Cross-references: EMBL:M74087  
 R:Worliarity, D.M.; Barringer, K.J.; Dodgson, J.B.; Richter, H.E.; Young, R.B.  
 DNA 6, 91-99, 1987  
 A:Title: Genomic clones encoding chicken myosin heavy-chain genes.  
 A:Reference number: A26821; MUID:87217964  
 A:Accession: A26321  
 A:Molecule type: DNA  
 A:Residues: 'F', 856-936 <MOR>  
 A:Cross-references: GB:M16557; NID:9212371; PIDN:AAA48970.1; PID:9212372  
 C:Genetics:  
 A:Introns: 886/3  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: actin binding; ATP; coiled coil; muscle; skeletal muscle

Query Match 93.5%; Score 29; DB 2; Length 936;  
 Best Local Similarity 85.7%; Pred. NO. 72;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7

Db 232 LASNMES 238

RESULT 15

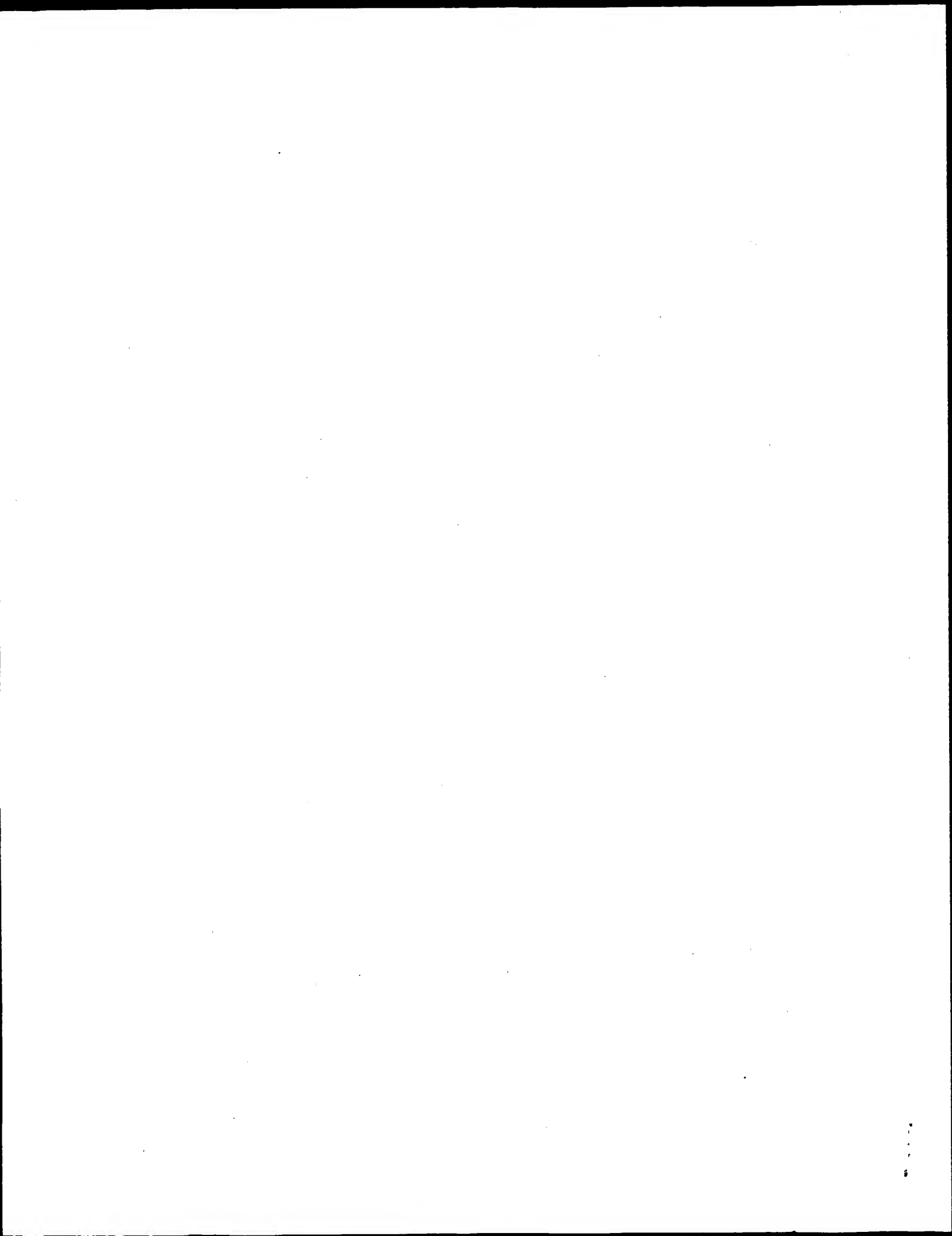
S24348  
 myosin heavy chain, embryonic and adult skeletal muscle (clone Cemb2) - chicken (fragment  
 C:Species: Gallus gallus (chicken)  
 C:Date: 03-Feb-1994 #sequence\_revision 06-Sep-1996 #text\_change 13-Feb-1998  
 C:Accession: S24348  
 R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.  
 J. Mol. Biol. 225, 1143-1151, 1992  
 A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isoform  
 A:Reference number: S24348; MUID:92309413  
 A:Accession: S24348  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-955 <MOO>  
 A:Cross-references: EMBL:M74085  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: ATP; coiled coil; muscle contraction; skeletal muscle; thick filament

Query Match 93.5%; Score 29; DB 2; Length 955;  
 Best Local Similarity 85.7%; Pred. NO. 73;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7

Db 250 LASNMES 256

Search completed: June 28, 2001, 15:58:46  
 Job time: 374 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:54:39 ; Search time 105.36 Seconds  
(without alignments)  
2.276 Million cell updates/sec

Title: US-09-724-406-30  
Perfect score: 31  
Sequence: 1 LASNLES 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	111	1 KV3J_MOUSE	P01662 mus musculus
2	31	100.0	111	1 KV3K_MOUSE	P01663 mus musculus
3	31	100.0	111	1 KV3R_MOUSE	P01670 mus musculus
4	31	100.0	111	1 KV3S_MOUSE	P01671 mus musculus
5	31	100.0	111	1 KV3T_MOUSE	P01672 mus musculus
6	31	100.0	131	1 KV3L_MOUSE	P01661 mus musculus
7	29	93.5	230	1 CUT1_FUSSC	Q99174 fusarium so
8	29	93.5	230	1 CUT1_FUSSO	P00590 fusarium so
9	29	93.5	1938	1 MYSS_CHICK	P13538 gallus gall
10	29	93.5	1940	1 MYSE_CHICK	P02565 gallus gall
11	28	90.3	186	1 NADM_PVRAB	Q9uyd4 pyrococcus
12	28	90.3	186	1 NADM_PVRHO	O58211 pyrococcus
13	28	90.3	376	1 KYVA_CAEEL	Q17963 caenorhabdi
14	28	90.3	1116	1 IN94_YEAST	P53751 saccharomyc
15	28	90.3	1225	1 Y309_MYCGE	P47551 mycoplasma
16	28	90.3	1549	1 YIR3_YEAST	P40438 saccharomyc
17	28	90.3	1549	1 YJW2_YEAST	P40890 saccharomyc
18	27	87.1	108	1 KVLB_HUMAN	P01594 homo sapien
19	27	87.1	108	1 KV3V_MOUSE	P01674 mus musculus
20	27	87.1	110	1 KV3P_MOUSE	P01668 mus musculus
21	27	87.1	111	1 KV3H_MOUSE	P01660 mus musculus
22	27	87.1	111	1 KV3L_MOUSE	P01664 mus musculus
23	27	87.1	111	1 KV3M_MOUSE	P01665 mus musculus
24	27	87.1	111	1 KV3N_MOUSE	P01666 mus musculus
25	27	87.1	111	1 KV3O_MOUSE	P01667 mus musculus
26	27	87.1	111	1 KV3Q_MOUSE	P01669 mus musculus
27	27	87.1	155	1 Y359_RICPR	Q9zdg9 rickettsia
28	27	87.1	323	1 RFC4_YEAST	P40339 saccharomyc
29	27	87.1	463	1 YIEQ_HAEIN	P44903 haemophilus
30	27	87.1	714	1 YFE7_YEAST	P43556 saccharomyc
31	27	87.1	1325	1 Y309_MYCPN	P75334 mycoplasma
32	26	83.9	111	1 KV3U_MOUSE	P01673 mus musculus
33	26	83.9	112	1 KV3G_MOUSE	P01659 mus musculus

34	26	83.9	132	1 KV3F_MOUSE	P01658 mus musculus
35	26	83.9	257	1 LSG4_HAEIN	P71398 haemophilus
36	26	83.9	258	1 Y875_SYNY3	P73555 synechocyst
37	26	83.9	648	1 KAPC_DICDI	P34099 dictyostell
38	26	83.9	796	1 MBN_DROME	P52302 drosophila
39	26	83.9	876	1 MYSS_HUMAN	P12882 homo sapien
40	25	80.6	146	1 RS18_YEAST	P35271 saccharomyc
41	25	80.6	175	1 CSF3_CANFA	P35834 canis famli
42	25	80.6	191	1 YDA4_SCHPO	Q10346 schizosacch
43	25	80.6	193	1 VIN3_BPT4	P13302 bacterioph
44	25	80.6	194	1 CSF3_FELCA	Q02708 felis silve
45	25	80.6	197	1 YQOC_CAEEL	Q09301 caenorhabdi

## ALIGNMENTS

```

RESULT 1
KV3J_MOUSE
ID KV3J_MOUSE STANDARD; PRT: 111 AA.
AC P01662;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG KAPPA CHAIN V-III REGION ABPC 22/PC 9245.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE (ABPC 22).
RX MEDLINE=79012520; PubMed=99744;
Mckean D.J., Bell M., Potter M.;
*Mechanisms of antibody diversity: multiple genes encode structurally
RT related mouse kappa variable regions.*;
RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).
RN [2]
RP SEQUENCE (PC 9245).
RX MEDLINE=79073152; PubMed=103003;
Weigert M., Garmaitan L., Loh E., Schilling J., Hood L.E.;
*Rearrangement of genetic information may produce immunoglobulin
RT diversity.*;
RL Nature 276:785-790(1978).
CC -1- MISCELLANEOUS: THE ABPC22 AND PC9241 SEQUENCES ARE IDENTICAL.
DR PIR: A01935; KMSM6.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 29 53
FT DOMAIN 39 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12041 MW; D7DF0609303453CE CRC64;

```

```

Query Match 100.0%; Score 31; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
QY 1 LASNLES 7
Db 54 LASNLES 60
RESULT 2
KV3K_MOUSE
ID KV3K_MOUSE STANDARD; PRT: 111 AA.
AC P01663;

```

21-JUL-1986 (Rel. 01, Created)  
21-JUL-1986 (Rel. 01, Last sequence update)  
15-JUL-1999 (Rel. 38, Last annotation update)  
IG KAPPA CHAIN V-III REGION PC 4050.  
Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE.  
MEDLINE=79073152; PubMed=103003;  
Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
"Rearrangement of genetic information may produce immunoglobulin  
diversity.";  
PIR: A01935; KWSM6.  
Nature 276:785-790(1978).  
InterPro: IPR003006; -.  
Pfam: PF00047; Ig; 1.  
IG KAPPA CHAIN V-III REGION PC 6684.  
Immunoglobulin V region.  
DOMAIN 1 23  
DOMAIN 24 38  
DOMAIN 39 53  
DOMAIN 54 60  
DOMAIN 61 92  
DOMAIN 93 101  
DOMAIN 102 111  
DISULFID 23 92  
NON\_TER 111 111  
SEQUENCE 111 AA; 12005 MW; 39DB7619313453CB CRC64;

Query Match 100.0%; Score 31; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.78;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
| | | | |  
DB 54 LASNLES 60

RESULT 3  
KV3R\_MOUSE  
ID KV3R\_MOUSE STANDARD; PRT; 111 AA.  
AC P01670;  
21-JUL-1986 (Rel. 01, Created)  
21-JUL-1986 (Rel. 01, Last sequence update)  
15-JUL-1999 (Rel. 38, Last annotation update)  
IG KAPPA CHAIN V-III REGION PC 6684.  
Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE.  
MEDLINE=79073152; PubMed=103003;  
Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
"Rearrangement of genetic information may produce immunoglobulin  
diversity.";  
PIR: A01938; KWSM84.  
Nature 276:785-790(1978).  
InterPro: IPR003006; -.  
Pfam: PF00047; Ig; 1.  
IG KAPPA CHAIN V-III REGION PC 6684.  
Immunoglobulin V region.  
DOMAIN 1 23  
DOMAIN 24 38  
DOMAIN 39 53  
DOMAIN 54 60  
DOMAIN 61 92  
DOMAIN 93 101  
DOMAIN 102 111  
DISULFID 23 92  
NON\_TER 111 111  
SEQUENCE 111 AA; 12039 MW; 1E4698834185826 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.78;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
| | | | |  
DB 54 LASNLES 60

RESULT 4  
KV3S\_MOUSE  
ID KV3S\_MOUSE STANDARD; PRT; 111 AA.  
AC P01671;  
21-JUL-1986 (Rel. 01, Created)  
21-JUL-1986 (Rel. 01, Last sequence update)  
15-JUL-1999 (Rel. 38, Last annotation update)  
IG KAPPA CHAIN V-III REGION PC 7175.  
Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE.  
MEDLINE=79073152; PubMed=103003;  
Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
"Rearrangement of genetic information may produce immunoglobulin  
diversity.";  
PIR: B01938; KWS75.  
Nature 276:785-790(1978).  
InterPro: IPR003006; -.  
Pfam: PF00047; Ig; 1.  
Immunoglobulin V region.  
DOMAIN 1 23  
DOMAIN 24 38  
DOMAIN 39 53  
DOMAIN 54 60  
DOMAIN 61 92  
DOMAIN 93 101  
DOMAIN 102 111  
DISULFID 23 92  
NON\_TER 111 111  
SEQUENCE 111 AA; 12010 MW; F041E89AA785823 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.78;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
| | | | |  
DB 54 LASNLES 60

RESULT 5  
KV3T\_MOUSE  
ID KV3T\_MOUSE STANDARD; PRT; 111 AA.  
AC P01672;  
21-JUL-1986 (Rel. 01, Created)  
21-JUL-1986 (Rel. 01, Last sequence update)  
15-JUL-1999 (Rel. 38, Last annotation update)  
IG KAPPA CHAIN V-III REGION PC 7940.  
Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE.  
MEDLINE=79073152; PubMed=103003;  
Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;  
"Rearrangement of genetic information may produce immunoglobulin  
diversity.";

```

FT DOMAIN          122   131      FRAMEWORK 4.
FT DISULFID       43    112      BY SIMILARITY.
FT NON_TER        131     .
SQ SEQUENCE       131 AA; 14291 MW; D212EC9F08DC880A CRC64;

Query Match              100.0%; Score 31; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LASNLES 7
    | | | | |
IDB 74 LASNLS 80

RESULT 7
CUTL_FUSSC
IID CUTL_FUSSC STANDARD; PRT: 230 AA.
Q99174;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
CUTINASE PRECURSOR (EC 3.1.1.-).
CUFA.
Fusarium solani (subsp. cucurbitae) (Nectria ipomoeae).
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hydrocales; Nectriaceae; Nectria.
NCBI_TaxID=57162;
[1]
SEQUENCE FROM N.A.
STRAIN=PGB 153;
MEDLINE=97254998; PubMed=9100380;
Crowhurst R.N., Binnie S.J., Bowen J.K., Hawthorne B.T., Plummer K.M.,
Rees-George J., Rikverink E.H., Templeton M.D.;
"Effect of disruption of a cutinase gene (cutA) on virulence and
tissue specificity of Fusarium solani f. sp. cucurbitae race 2 toward
Cucurbita maxima and C. moschata.";
Mol. Plant Microbe Interact. 10:355-368(1997).
-!- FUNCTION: ALLOWS PATHOGENIC FUNGI TO PENETRATE THROUGH THE
CUTICULAR BARRIER INTO THE HOST PLANT DURING THE INITIAL STAGE
OF THE FUNGAL INFECTION (BY SIMILARITY).
-!- CATALYTIC ACTIVITY: HYDROLYSIS OF CUTIN, A POLYESTER THAT
FORMS THE STRUCTURE OF PLANT CUTICLE.
-!- SUBCELLULAR LOCATION: SECRETED.
-!- PM: THE 2 DISULFIDE BONDS PLAY A CRITICAL ROLE IN HOLDING THE
CATALYTIC RESIDUES IN JUXTA-POSITION; REDUCTION OF THE DISULFIDE
BRIDGES RESULTS IN THE COMPLETE INACTIVATION OF THE ENZYME (BY
SIMILARITY).
-!- SIMILARITY: BELONGS TO THE CUTINASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U63335; AAA05922.1; -.
HSSP; P00590; 2CUT.
InterPro: IPR000675; -.
Pfam; PF01083; Cutinase; 1.
PRINTS; PR00129; CUTINASE.
PROSITE; PS00155; CUTINASE_1; 1.
PROSITE; PS00931; CUTINASE_2; 1.
Hydrolase; Serine esterase; Glycoprotein; Signal.
SIGNAL 1 16
CHAIN 17 230
BY SIMILARITY.
CUTINASE.
FT CHAIN 17 230
FT DISULFID 47 194
BY SIMILARITY.
FT DISULFID 125 187
BY SIMILARITY.
FT ACT_SITE 136 136
BY SIMILARITY.
FT ACT_SITE 191 191
BY SIMILARITY.
FT ACT_SITE 204 204
BY SIMILARITY.

```

SQ SEQUENCE 230 AA; 23902 MW; 05FB3C3326405AA CRC64;  
 Query Match 93.5%; Score 29; DB 1; Length 230;  
 Best Local Similarity 85.7%; Pred. No. 5.7;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LASNLES 7  
 Db 71 IASNLES 77  
 RESULT 8  
 ID CUTI\_FUSSO STANDARD; PRT; 230 AA.  
 AC P00590; 1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CUTINASE PRECURSOR (EC 3.1.1.-).  
 GN CUTA.  
 OS Fusarium solani (subsp. pisi) (Nectria haematococca).  
 OC Eukaryota; Fungi; Ascomycota; Mitosporic Ascomycota; Fusarium.  
 OX NCBI\_TaxID=109625;  
 RN (1)  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN-T-8;  
 RA Soliday C.L., Flurkey W.H., Okita T.W., Kolattukudy P.E.;  
 RT "Cloning and structure determination of cDNA for cutinase, an enzyme  
 involved in fungal penetration of plants.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3939-3943(1984).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89197761; PubMed=2703464;  
 RA Soliday C.L., Dickman M.B., Kolattukudy P.E.;  
 RT "Structure of the cutinase gene and detection of promoter activity in  
 the 5'-flanking region by fungal transformation.";  
 RL J. Bacteriol. 171:1942-1951(1989).  
 RN (3)  
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
 RX MEDLINE=92220194; PubMed=1560844;  
 RA Martinez C., de Geus P., Lauwers M., Matthysens G., Cambillau C.;  
 RT "Fusarium solani cutinase is a lipolytic enzyme with a catalytic  
 serine accessible to solvent.";  
 RL Nature 356:615-618(1992).  
 RN (4)  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=94114517; PubMed=8286366;  
 RA Martinez C., Nicolas A., van Tilbeurgh H., Egloff M.-P., Cudrey C.,  
 Verger R., Cambillau C.;  
 RT "Cutinase, a lipolytic enzyme with a preformed oxyanion hole.";  
 RL Biochemistry 33:83-89(1994).  
 RN (5)  
 RP X-RAY CRYSTALLOGRAPHY (1.0 ANGSTROMS).  
 RX MEDLINE=97318923; PubMed=9175860;  
 RA Longhi S., Czjzek M., Lamzin V., Nicolas A., Cambillau C.;  
 RT "Atomic resolution (1.0 Å) crystal structure of Fusarium solani  
 cutinase: stereochemical analysis.";  
 RL J. Mol. Biol. 268:779-799(1997).  
 RN (6)  
 RP X-RAY CRYSTALLOGRAPHY (1.15 ANGSTROMS) OF 33-230.  
 RA Nicolas A., Martinez C., Cambillau C.;  
 RL Submitted (MAR-1997) to the PDB data bank.  
 RN (7)  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=98046750; PubMed=9385640;  
 RA Prompers J.J., Groenewegen A., van Schaik R.C., Pepermans H.A.M.,  
 Hilbers C.W.;  
 RT "1H, 13C, and 15N resonance assignments of Fusarium solani pisi  
 cutinase and preliminary features of the structure in solution.";  
 RL Protein Sci. 6:2375-2384(1997).  
 CC -1- FUNCTION: ALLOWS PATHOGENIC FUNGI TO PENETRATE THROUGH THE  
 CUTICULAR BARRIER INTO THE HOST PLANT DURING THE INITIAL STAGE

OF THE FUNGAL INFECTION.  
 -1- CATALYTIC ACTIVITY: HYDROLYSIS OF CUTIN, A POLYESTER THAT  
 FORMS THE STRUCTURE OF PLANT CUTICLE.  
 -1- SUBCELLULAR LOCATION: SECRETED.  
 -1- INDUCTION: AFTER CONTACT WITH CUTIN  
 -1- PTM: THE 2 DISULFIDE BONDS PLAY A CRITICAL ROLE IN HOLDING THE  
 CATALYTIC RESIDUES IN JUXTA-POSITION; REDUCTION OF THE DISULFIDE  
 BRIDGES RESULTS IN THE COMPLETE INACTIVATION OF THE ENZYME.  
 -1- SIMILARITY: BELONGS TO THE CUTINASE FAMILY.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; K02640; AAA33334.1; -  
 EMBL; M29759; AAA33335.1; -  
 PIR; A00731; UVFUS.  
 PIR; A32836; A32836.  
 PDB; 1CUS; 31-JUL-94.  
 PDB; 2CUT; 31-AUG-94.  
 PDB; 1CVA; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUF; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUH; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUU; 11-JUL-96.  
 PDB; 1CIV; 11-JUL-96.  
 PDB; 1COW; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUG; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUB; 11-JUL-96.  
 PDB; 1CUC; 11-JUL-96.  
 PDB; 1CUD; 11-JUL-96.  
 PDB; 1CUE; 11-JUL-96.  
 PDB; 1CUI; 11-JUL-96.  
 PDB; 1CUJ; 11-JUL-96.  
 PDB; 1CUX; 11-JUL-96.

FT ACT\_SITE 136 136  
 FT ACT\_SITE 191 191  
 FT ACT\_SITE 204 204  
 FT CONFLICT 48 48 R -> A (IN REF. 2).  
 FT CONFLICT 94 94 R -> A (IN REF. 2).  
 SQ SEQUENCE 230 AA; 23982 MW; 7253ACAA657AD1AB CRC64;

Query Match 93.5%; Score 29; DB 1; Length 230;  
 Best Local Similarity 85.7%; Pred. No. 5.7;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASNLES 7  
 Db 71 IASNLES 77

RESULT 9  
 MYSS\_CHICK STANDARD; PRT; 1938 AA.  
 AC P13538;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE MYOSIN HEAVY CHAIN, SKELETAL MUSCLE, ADULT.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE OF 1-205.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041767; PubMed=1939027;  
 RA Hayashida M., Maita T., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: I.  
 RL Sequence of the amino-terminal 23 kDa fragment.";  
 RJ J. Biochem. 110:54-59(1991).  
 RN [2]  
 RP SEQUENCE OF 206-636.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041768; PubMed=1939028;  
 RA Komine Y., Maita T., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: II.  
 RL Sequence of the 50 kDa fragment of subfragment-1.";  
 RJ J. Biochem. 110:60-67(1991).  
 RN [3]  
 RP SEQUENCE OF 637-837.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041769; PubMed=1939029;  
 RA Maita T., Miyaniishi T., Matsuzono K., Tanioka Y., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: III.  
 RL Sequence of the 22 kDa fragment and the alignment of the 23 kDa, 50  
 kDa, and 22 kDa fragments.";  
 RJ J. Biochem. 110:68-74(1991).  
 RN [4]  
 RP SEQUENCE OF 838-1938.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041770; PubMed=1939030;  
 RA Maita T., Yajima E., Nagata S., Miyaniishi T., Nakayama S., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: IV.  
 RL Sequence of the rod, and the complete 1,938-residue sequence of the  
 heavy chain.";  
 RJ J. Biochem. 110:75-87(1991).  
 RN [5]  
 RP PRELIMINARY SEQUENCE OF 1-808.  
 RX MEDLINE=87092420; PubMed=3467365;  
 RA Maita T., Hayashida M., Tanioka Y., Komine Y., Matsuda G.;  
 RT "The primary structure of the myosin head.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:416-420(1987).  
 RN [6]  
 RP SEQUENCE OF 842-1270.  
 RX MEDLINE=90121764; PubMed=2610940;

RA Watanabe B.;  
 RT "Complete amino-acid sequence of subfragment-2 in adult chicken  
 skeletal muscle myosin.";  
 RL Biol. Chem. Hoppe-Seyler 370:1027-1034(1989).  
 RN [7]  
 RP SEQUENCE OF 852-1108.  
 RX MEDLINE=89374803; PubMed=2775482;  
 RA Watanabe B.;  
 RT "Amino-acid sequence of the short subfragment-2 in adult chicken  
 skeletal muscle myosin.";  
 RL Biol. Chem. Hoppe-Seyler 370:549-558(1989).  
 RN [8]  
 RP SEQUENCE OF 1145-1270.  
 RX MEDLINE=89228549; PubMed=2713098;  
 RA Watanabe B.;  
 RT "Amino-acid sequence of the hinge region in chicken myosin  
 subfragment-2.";  
 RL Biol. Chem. Hoppe-Seyler 370:55-61(1989).  
 RN [9]  
 RP SEQUENCE OF 1857-1938 FROM N.A.  
 RX MEDLINE=87217964; PubMed=3034534;  
 RA Moriarty D.M., Barringer K.J., Dodgson J.B., Richter H.E.,  
 Young R.B.;  
 RT "Genomic clones encoding chicken myosin heavy-chain genes.";  
 RL DNA 6:91-99(1987).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-843.  
 RX MEDLINE=93303624; PubMed=8316857;  
 RA Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R.,  
 Tomchick D.R., Benning M.M., Winkelman D.A., Wesenberg G.,  
 Holden H.M.;  
 RT "Three-dimensional structure of myosin subfragment-1: a molecular  
 motor.";  
 RL Science 261:50-58(1993).  
 CC -I- FUNCTION: MUSCLE CONTRACTION. MYOSIN IS A PROTEIN THAT BINDS TO  
 P-ACTIN AND HAS ATPASE ACTIVITY THAT IS ACTIVATED BY P-ACTIN.  
 CC -I- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -I- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -I- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -I- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HWM). IT CAN LATER BE  
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 SUBFRAGMENT (S2).  
 CC -I- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,  
 WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE  
 CONSERVED.

-----  
 THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: M16557; AAA48970.1;  
 PIR: PX0050; PX0051.  
 PIR: A26821; A26821.  
 PIR: S02082; S02082.  
 PIR: S04501; S04501.  
 PIR: S05515; S05515.  
 PDB: 2MYS; 1I-JAN-97.  
 InterPro: IPR000048;  
 InterPro: IPR001609;  
 InterPro: IPR002928;  
 Pfam: PF00612; IQ: 1.  
 Pfam: PF01576; Myosin\_tail; 1.





RC STRAIN=ORSAY;  
 RA Hellig R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome  
 structure and evolution.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + NICOTINAMIDE RIBONUCLEOTIDE =  
 CC DIPHOSPHATE + NAD(+).  
 CC -1- PATHWAY: NAD BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEL NMN ADENYLTRANSFERASE  
 CC FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: AJ248288; CAB50478.1; -;  
 CC InterPro: IPR001994; -;  
 CC Pfam: PF01467; Cytidylyltransf; 1.  
 CC Transferease: Nucleotidyltransferase; NAD.  
 CC SEQUENCE 186 AA; 21417 MW; E8230B688481386E CRC64;

Query Match 90.3%; Score 28; DB 1; Length 186;  
 Best Local Similarity 85.7%; Pred. No. 7.9;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 II:IIII  
 DB 166 LATNLES 172

RESULT 12  
 NADM\_PYRHO STANDARD; PRT; 186 AA.  
 AC O58211;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE NICOTINAMIDE-NUCLEOTIDE ADENYLTRANSFERASE (EC 2.7.7.1) (NAD(+)  
 DE PYROPHOSPHORYLASE) (NAD(+)) DIPHOSPHORYLASE (NMN ADENYLTRANSFERASE).  
 GN PH0464.  
 OS Pyrococcus horikoshii.  
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 CC NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3.  
 RX MEDLINE=98344137; PubMed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -1- CATALYTIC ACTIVITY: ATP + NICOTINAMIDE RIBONUCLEOTIDE =  
 CC DIPHOSPHATE + NAD(+).  
 CC -1- PATHWAY: NAD BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEL NMN ADENYLTRANSFERASE  
 CC FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: AF000002; BAA29550.1; -;  
 CC InterPro: IPR001994; -;  
 CC Pfam: PF01467; Cytidylyltransf; 1.  
 CC Transferease: Nucleotidyltransferase; NAD.  
 CC SEQUENCE 186 AA; 21392 MW; 1CE5A40C884ADB34 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 186;  
 Best Local Similarity 85.7%; Pred. No. 7.9;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 II:IIII  
 DB 166 LATNLES 172

RESULT 13  
 YKY4\_CAEEL STANDARD; PRT; 376 AA.  
 AC Q17963;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL 40.4 KDA TRP-ASP REPEATS CONTAINING PROTEIN C14B1.4 IN  
 DE CHROMOSOME III.  
 GN C14B1.4.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 CC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Harris B.;

RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO K04G11.4 AND ZC302.2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: Z37139; CAA85487.1; -;  
 CC WormPep: C14B1.4; CE00901.  
 CC InterPro: IPR001680; -;  
 CC Pfam: PF00400; WD40; 7.  
 CC PRINTS: PR00320; GPROTEINRPT  
 CC PROSITE: PS00678; WD\_REPEATS\_1; 5.  
 CC PROSITE: PS50082; WD\_REPEATS\_2; 6.  
 CC PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 CC Hypothetical protein; Repeat; WD repeat.  
 KW REPEAT 85 115 WD 1.  
 FT REPEAT 127 157 WD 2.  
 FT REPEAT 169 199 WD 3.  
 FT REPEAT 211 241 WD 4.  
 FT REPEAT 254 284 WD 5.  
 FT REPEAT 296 329 WD 6.  
 FT REPEAT 341 373 WD 7.  
 SQ SEQUENCE 376 AA; 40393 MW; 3204DC36FE58FC19 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 376;  
 Best Local Similarity 85.7%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 LASNLES 7
Db 272 LASNLS 278

RESULT 14
YN94_YEAST
ID_YN94_YEAST STANDARD; PRT; 1116 AA.
AC P53751;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PUTATIVE 125.2 KDA MEMBRANE GLYCOPROTEIN IN BIO3-HXT17 INTERGENIC
DE REGION.
GN YNR065C OR N3539.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Duesterhoeft A., Floeth M., Fritz C., Heuss-Neitzel D.,
RA Hilbert H., Moestl D.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE PEPI FAMILY OF MEMBRANE GLYCOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z71680; CAA96347.1; -
CC SGD; S0005348; YNR065C.
CC InterPro; IPR002860; -
CC Pfam; PF02012; BNR; 7.
KW Hypothetical protein; Transmembrane; Glycoprotein.
FT TRANSMEM 934 957
FT CARBOHYD 35 35 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 846 846 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 985 985 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 1116 AA; 125199 MW; C0361878FE4DAB90 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 1116;
Best Local Similarity 85.7%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LASNLES 7
Db 534 LATNLES 540

RESULT 15
Y309_MYCGE
ID Y309_MYCGE STANDARD; PRT; 1225 AA.
AC P47551; O49317;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL LIPOPROTEIN MG309 PRECURSOR.
GN MG309.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AJCC 33530 / G-37;

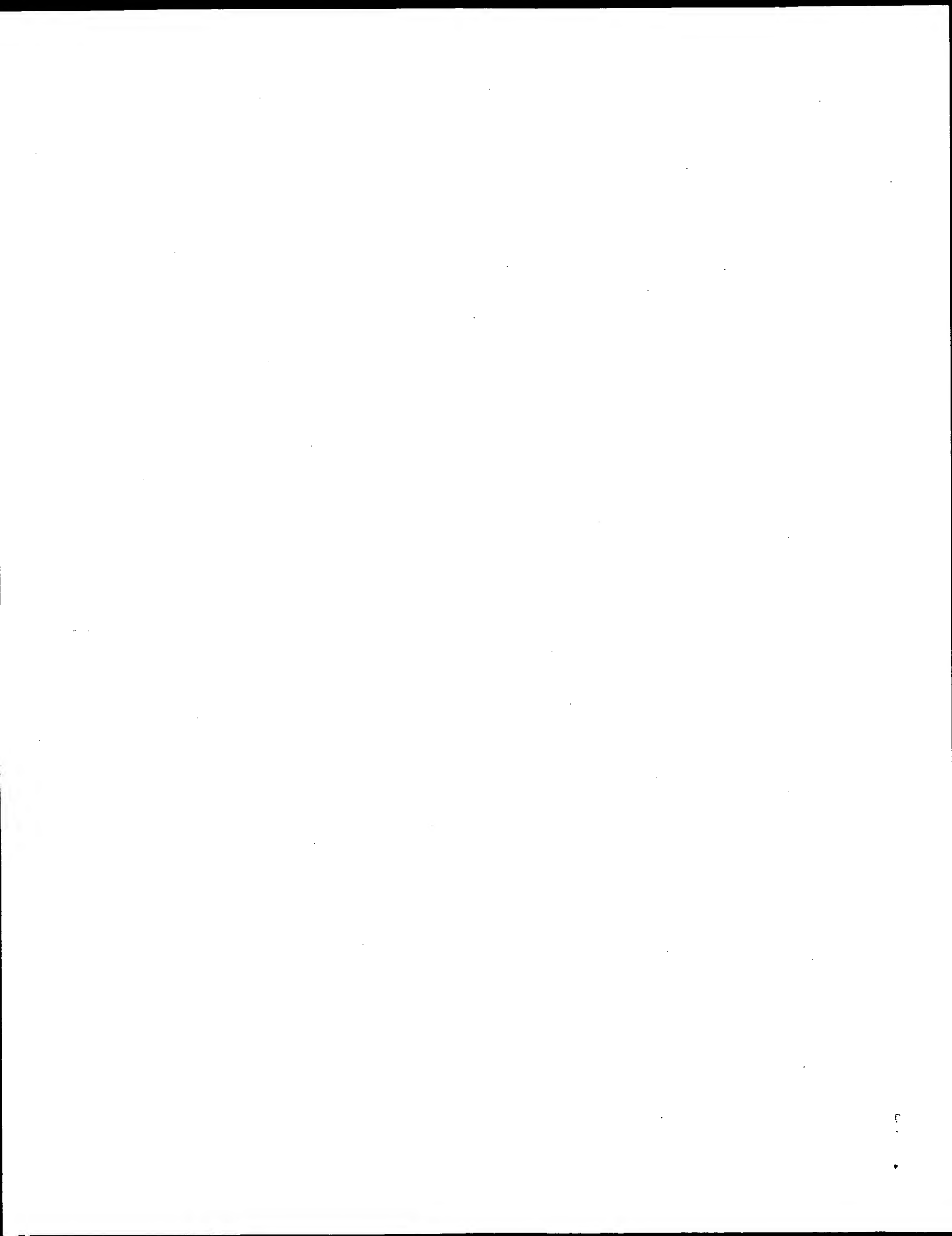
MEDLINE=96026346; PubMed=7569993;
RX Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 1138-1224 FROM N.A.
RC STRAIN-AJCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE MG307 / MG309 / MG338 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39711; AAC71531.1; -
CC EMBL; U02200; AAD12488.1; -
CC TIGR; MG309; -
KW Hypothetical protein; Lipoprotein; Membrane; Signal.
FT SIGNAL 1 27
FT CHAIN 28 1225 HYPOTHETICAL LIPOPROTEIN MG309.
FT LIPID 28 28 N-ACYL DIGLYCERIDE (POTENTIAL).
FT CONFLICT 1185 1185 L -> V (IN REF. 2).
SQ SEQUENCE 1225 AA; 138375 MW; C3E4BF5B4319B6E8 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 1225;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LASNLES 7
Db 286 LASNLS 292

Search completed: June 28, 2001, 15:54:39
Job time: 127 sec
```





100

---

[illegible]

1

QY 1 LASNLES 7

Db 104 LASNLES 110

RESULT 2

Q9VHD5 PRELIMINARY; PRT; 826 AA.

AC Q9VHD5;

DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE MYOSIN HEAVY CHAIN (FRAGMENT).

GN MHC-4.

OS Rana catesbeiana (Bull frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.

OX NCBI\_TaxID=8400;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TAIL MUSCLE; HINDLIMB MUSCLE;

RA Hu H., Merrifield P., Atkinson B.G.;

RT "Expression of the Myosin Heavy Chain Genes in the Tail Muscle of

RT Thyroid Hormone-induced Metamorphosing Rana catesbeiana Tadpoles.";

RL Dev. Genet. 0:0-0(1999).

DR EMBL: AF097907; AAD13772.1; -.

DR InterPro: IPR000533; -.

DR InterPro: IPR002928; -.

DR Pfam: PF01576; Myosin\_tail; 1.

DR PRINTS: PR00194; TROPOMYOSIN.

FT NON\_TER 1

SQ SEQUENCE 826 AA; 95361 MW; 8F9E3B0579D0F483 CRC64;

Query Match

Best Local Similarity 100.0%; Score 31; DB 13; Length 826;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7

Db 122 LASNLES 128

RESULT 3

Q9V706 PRELIMINARY; PRT; 856 AA.

AC Q9V706;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE CGI3942 PROTEIN.

GN CGI3942.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY.

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003815; AAF58260.1; -.

DR FlyBase: FBgn003922; CG13942.

DR InterPro: IPR002965; -.

DR PRINTS: PR01217; PRICHEXTENSIN.

SQ SEQUENCE 856 AA; 93777 MW; 3B6C3F80DFEA12B7 CRC64;

Query Match 93.5%; Score 29; DB 5; Length 856;

Best Local Similarity 85.7%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7

Db 845 IASNLES 851

RESULT 4

O13228 PRELIMINARY; PRT; 1939 AA.

AC O13228;

DT 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE MYOSIN HEAVY CHAIN.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=WHITE LEHORN; TISSUE=PECTORALIS MUSCLE;

RA Chao T.H., Bandman E., Moore L.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: U87231; AAB47555.1; -.

DR HSSP: P08799; IMMD.

DR InterPro: IPR000048; -.

DR InterPro: IPR001609; -.

DR InterPro: IPR002928; -.

DR Pfam: PF00063; myosin\_head; 1.

DR Pfam: PF00612; IQ; 1.

DR Pfam: PF01576; Myosin\_tail; 1.

DR PRINTS: PR00193; MYOSIN\_HEAVY.

DR ProDom: PD000355; -; 1.

DR SMART: SM00015; IQ; 1.

SQ SEQUENCE 1939 AA; 223178 MW; 4D4E3E037B7550E9 CRC64;

Query Match 93.5%; Score 29; DB 13; Length 1939;  
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 ||||:||  
 Db 1234 LASNLES 1240

RESULT 5  
 Q9PTY2 PRELIMINARY; PRT; 1939 AA.  
 AC Q9PTY2  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE FAST MYOSIN HEAVY CHAIN.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEART;  
 RA Machida S., Matsuoaka R., Noda S., Hiratsuka E., Takagaki Y., Oana S.,  
 Furutani Y., Nakajima H., Takao A., Momma K.;  
 RT "Evidence for the expression of neonatal skeletal myosin heavy chain  
 in primary myocardium and cardiac conduction tissue in the developing  
 chick heart."  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB021180; BAA82233.1; -;  
 DR HSSP: P08799; ILVX.  
 DR InterPro: IPR000048; -;  
 DR InterPro: IPR000533; -;  
 DR InterPro: IPR001609; -;  
 DR InterPro: IPR002928; -;  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR PRINTS: PR00194; TROPOMYOSIN.  
 DR SMART: SM00242; MYSC; 1.  
 SQ SEQUENCE 1939 AA; 222940 MW; 0AAC40D61A2D07F3 CRC64;

Query Match 93.5%; Score 29; DB 13; Length 1939;  
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 ||||:||  
 Db 1235 LASNLES 1241

RESULT 6  
 Q9DGM4 PRELIMINARY; PRT; 1941 AA.  
 AC Q9DGM4  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE FAST MYOSIN HEAVY CHAIN ISOFORM 3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Q., Bandman E.;

Query Match 93.5%; Score 29; DB 13; Length 1944;  
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 ||||:||  
 Db 1239 LASNLES 1245

RESULT 8  
 Q22481 PRELIMINARY; PRT; 367 AA.  
 AC Q22481  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE SIMILARITY TO C. ELEGANS HYPOPHETICAL PROTEIN.  
 GN T1484.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Willson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
 Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,

Query Match 93.5%; Score 29; DB 13; Length 1941;  
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 ||||:||  
 Db 1236 LASNLES 1242

RESULT 7  
 Q9DGM5 PRELIMINARY; PRT; 1944 AA.  
 AC Q9DGM5  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE FAST MYOSIN HEAVY CHAIN ISOFORM 2.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Q., Bandman E.;

Query Match 93.5%; Score 29; DB 13; Length 1944;  
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 ||||:||  
 Db 1239 LASNLES 1245

RESULT 8  
 Q22481 PRELIMINARY; PRT; 367 AA.  
 AC Q22481  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE SIMILARITY TO C. ELEGANS HYPOPHETICAL PROTEIN.  
 GN T1484.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Willson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
 Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Favello T.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U50191; AAA91241.1; -;  
 DR InterPro; IPR003406; -;  
 DR Pfam; PF02485; Branch; 1.  
 SQ SEQUENCE 367 AA; 43010 MW; 72CB4A155DD447D4 CRC64;

Query Match 90.3%; Score 28; DB 5; Length 367;  
 Best Local Similarity 85.7%; Pred. No. 88;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 Db 76 LSSNLES 82

RESULT 9  
 Q9X5A4 PRELIMINARY; PRT; 393 AA.  
 AC Q9X5A4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE FLAGELLAR SWITCH PROTEIN FLY.  
 GN FLY.  
 OS Treponema denticola.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=158;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35405;  
 RA Stamm L.V., Bergen H.L.;  
 RT "Sequences of the tap1, flgD, flgE, orf4, motA, motB, flilL, flilM,  
 RT flilY, and flilP Treponema denticola genes.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF122909; AAD20623.1; -;  
 DR InterPro; IPR001172; -;  
 DR InterPro; IPR001543; -;  
 DR Pfam; PF01052; Spoa; 1.  
 DR PRINTS; PR00956; FLGMOTORFLIN.  
 KW Flagella.  
 SQ SEQUENCE 393 AA; 41701 MW; EF9B924DCC2AA342 CRC64;

Query Match 90.3%; Score 28; DB 2; Length 393;  
 Best Local Similarity 85.7%; Pred. No. 94;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 Db 54 LSSNLES 60

RESULT 10  
 O82521 PRELIMINARY; PRT; 459 AA.  
 AC O82521;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE PUTATIVE AMINOTRANSFERASE.  
 OS Capsicum chinense.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; I;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;

OC Solanales; Solanaceae; Capsicum.  
 OX NCBI\_TaxID=80379;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HABANERO;  
 RA Aluru M., Curry J., O'Connell M.A.;  
 RT "Nucleotide sequence of a probable aminotransferase from habanero  
 RT chile.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF095149; AAC78480.1; -;  
 DR HSPF; P04181; 20AT.  
 DR InterPro; IPR000954; -;  
 DR Pfam; PF0202; aminotran\_3; 1.  
 DR PROSITE; PS00600; AA\_TRANSFER\_CLASS\_3; UNKNOWN\_1.  
 KW Transferase; Aminotransferase.  
 SQ SEQUENCE 459 AA; 50729 MW; 02ABBAD728B524E4 CRC64;

Query Match 90.3%; Score 28; DB 10; Length 459;  
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 Db 200 LANNLES 206

RESULT 11  
 Q9YHD7 PRELIMINARY; PRT; 708 AA.  
 AC Q9YHD7;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE MYOSIN HEAVY CHAIN (FRAGMENT).  
 GN MHC-2.  
 OS Rana catesbeiana (Bull frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.  
 OX NCBI\_TaxID=8400;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-TAIL MUSCLE, HINDLIMB MUSCLE;  
 RA Hu H., Merrifield P., Atkinson B.G.;  
 RT "Expression of the Myosin Heavy Chain Genes in the Tail Muscle of  
 RT Thyroid Hormone-induced Metamorphosing Rana catesbeiana Tadpoles.";  
 RL Dev. Genet. 0:0-0(1999).  
 DR EMBL; AF097905; AAD13770.1; -;  
 DR InterPro; IPR000533; -;  
 DR InterPro; IPR002928; -;  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00194; TROPOMYOSIN.  
 FT NON\_TER  
 SQ SEQUENCE 708 AA; 81823 MW; BE2D01700017BC4C CRC64;

Query Match 90.3%; Score 28; DB 13; Length 708;  
 Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 Db 4 LASNLEN 10

RESULT 12  
 Q9YHD8 PRELIMINARY; PRT; 879 AA.  
 AC Q9YHD8;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE MYOSIN HEAVY CHAIN (FRAGMENT).



GN MHC-1.  
 OS Rana catesbeiana (Bull frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.  
 OX NCBI\_TaxID=8400;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TAIL MUSCLE, HINDLIMB MUSCLE;  
 RA Hu H., Merrifield P., Atkinson B.G.;  
 RT "Expression of the Myosin Heavy Chain Genes in the Tail Muscle of  
 RT Thyroid Hormone-Induced Metamorphosing Rana catesbeiana Tadpoles.";  
 RL Dev. Genet. 0:0-0(1999).  
 DR EMBL; AF097904; AAD13769.1; -.  
 DR InterPro; IPR000533; -.  
 DR pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00194; TROPOMYOSIN.  
 FT NON\_TER 1  
 SQ SEQUENCE 879 AA; 101710 MW; 1C456851E968A3D8 CRC64;

Query Match 90.3%; Score 28; DB 13; Length 879;  
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 Db 171 LASNLEN 177

RESULT 13  
 Q9FHH2 PRELIMINARY; PRT; 990 AA.  
 AC Q9FHH2;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 101 KDA HEAT SHOCK PROTEIN, HSP101-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=99397451; PubMed=10470850;  
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.  
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
 RT Pl and TAC clones.";  
 RL DNA Res. 6:183-195(1999).  
 DR EMBL; AB018118; BAB09589.1; -.  
 KW Heat shock.  
 SQ SEQUENCE 990 AA; 108709 MW; A0BA396F566466CB CRC64;

Query Match 90.3%; Score 28; DB 10; Length 990;  
 Best Local Similarity 85.7%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 ||:||||  
 Db 395 LANNLES 401

RESULT 14  
 Q95054 PRELIMINARY; PRT; 125 AA.  
 AC Q95054;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE IRON SUPEROXIDE DISMUTASE (EC 1.15.1.1) (FRAGMENT).  
 GN SOD2.  
 OS Trichomonas vaginalis.  
 OC Eukaryota; Parabasalidea; Trichomonadida; Trichomonas.  
 OX NCBI\_TaxID=5722;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30001;  
 RA Viscogliosi E., Durieux I., Delgado-Viscogliosi P., Bayle D., Dive D.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30001;  
 RA Touzel J.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30001;  
 RX MEDLINE=97047378; PubMed=8892298;  
 RA Viscogliosi E., Durieux I., Delgado-Viscogliosi P., Bayle D., Dive D.;  
 RT "Phylogenetic implication of iron-containing superoxide dismutase  
 RT genes from trichomonad species.";  
 RL Mol. Biochem. Parasitol. 80:209-214(1996).  
 CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).  
 CC -!- COFACTOR: COPPER AND ZINC OR IRON OR MANGANESE.  
 CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL; Z70671; CAA94528.1; -.  
 DR HSP; P09157; IISC.  
 DR InterPro; IPR001189; -.  
 DR Pfam; PF00081; sodfe; 1.  
 KW Oxidoreductase.  
 FT NON\_TER 1  
 FT NON\_TER 125  
 FT NON\_TER 125  
 SQ SEQUENCE 125 AA; 13826 MW; 108A018B073F60DD CRC64;

Query Match 87.1%; Score 27; DB 5; Length 125;  
 Best Local Similarity 85.7%; Pred. No. 51;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 |||||  
 Db 64 LASNFES 70

RESULT 15  
 Q95055 PRELIMINARY; PRT; 125 AA.  
 AC Q95055;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE IRON SUPEROXIDE DISMUTASE (EC 1.15.1.1) (FRAGMENT).  
 GN SOD3.  
 OS Trichomonas vaginalis.  
 OC Eukaryota; Parabasalidea; Trichomonadida; Trichomonas.  
 OX NCBI\_TaxID=5722;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30001;  
 RA Viscogliosi E., Durieux I., Delgado-Viscogliosi P., Bayle D., Dive D.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30001;  
 RA Touzel J.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 30001;  
 RX MEDLINE=97047378; PubMed=8892298;  
 RA Viscogliosi E., Durieux I., Delgado-Viscogliosi P., Bayle D., Dive D.;  
 RT "Phylogenetic implication of iron-containing superoxide dismutase  
 genes from trichomonad species."; Mol. Biochem. Parasitol. 80:209-214(1996).  
 RL  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: COPPER AND ZINC OR IRON OR MANGANESE.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 CC EMBL: Z70672; CAA94529.1; -.  
 DR HSP: P09157; IISC.  
 DR InterPro: IPR001189; -.  
 DR Pfam: PF00081; sodfe; 1.  
 KW Oxidoreductase.  
 FT NON\_TER 1 1  
 FT NON\_TER 125 125  
 SQ SEQUENCE 125 AA; 13776 MW; 12CCEBCB433B80DD CRC64;

Query Match 87.1%; Score 27; DB 5; Length 125;  
 Best Local Similarity 85.7%; Pred. No. 51;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LASNLES 7  
 Db 64 LASNFES 70

Search completed: June 28, 2001, 16:08:27  
 Job time: 955 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:14:39 ; Search time 362.28 Seconds  
(without alignments)  
1.506 Million cell updates/sec

Title: US-09-724-406-32  
Perfect score: 50  
Sequence: 1 QHSGELPFT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_0601.\*  
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	84.0	111	17	AA1985242 Humanised antibody
2	42	84.0	111	17	AA1985236 Murine 206 antibody
3	40	80.0	111	16	AA1985480 Mouse antibody lig
4	40	80.0	125	12	AA195440 Light chain variab
5	40	80.0	132	17	AA195440 Anti-human gp39 MA
6	39	78.0	405	21	AA195441 Human ORFX ORF2205
7	37	74.0	9	16	AA195475 Mouse antibody lig
8	37	74.0	21	15	AA195176 GM-CSF receptor al
9	37	74.0	111	13	AA195721 Humanised VL regio
10	37	74.0	111	22	AA195961 Murine Fd79 antibo
11	37	74.0	111	22	AA195962 Humanised Fd79 ant

12	37	74.0	218	16	AA1975457 Mouse antibody FB3
13	37	74.0	218	16	AA1975459 Mouse antibody F4-
14	36	72.0	9	21	AA1910000 H. pylori HSP60-bi
15	36	72.0	111	15	AA1960528 Light chain variab
16	36	72.0	111	21	AA1910019 H. pylori HSP60-bi
17	36	72.0	212	15	AA1952659 Porphyrin antibody
18	36	72.0	249	21	AA1956601 Human prostate can
19	36	72.0	523	21	AA190231 Zebrafish PTH3R re
20	36	72.0	640	20	AA192918 Human CHD1 protein
21	36	72.0	648	20	AA192917 Human CHD1 protein
22	35	70.0	307	21	AA193276 Amino acid sequenc
23	35	70.0	312	21	AA193298 DNA encoding a pol
24	35	70.0	307	21	AA193298 Human ORFX ORF49
25	35	70.0	539	21	AA191185 Human nucleic acid
26	35	70.0	782	21	AA192219 Human ORFX ORF1983
27	34	68.0	56	19	AA192687 Ikaro protein SEQ
28	34	68.0	56	19	AA192688 Ikaro protein SEQ
29	34	68.0	56	19	AA192689 Ikaro protein SEQ
30	34	68.0	84	21	AA192076 PRDI-BFL protein z
31	34	68.0	93	16	AA1959534 Anti-HCV antibody
32	34	68.0	112	21	AA1939077 Human secreted pro
33	34	68.0	136	21	AA1939073 Human secreted pro
34	34	68.0	137	21	AA1939071 Human secreted pro
35	34	68.0	179	21	AA1963193 Gene 8 human score
36	34	68.0	197	21	AA192630 Human ORFX ORF2394
37	34	68.0	197	21	AA194360 Human cell cycle r
38	34	68.0	219	21	AA193935 Human cancer assoc
39	34	68.0	252	17	AA1939364 Human REST protein
40	34	68.0	308	17	AA1939369 Human REST protein
41	34	68.0	318	21	AA193899 Amino acid sequenc
42	34	68.0	318	21	AA193900 Amino acid sequenc
43	34	68.0	334	17	AA192020 Ikaro protein. N
44	34	68.0	334	19	AA192677 Ikaro isoform. U
45	34	68.0	334	19	AA1970969 Human Ikaro prote

## ALIGNMENTS

RESULT 1

AA1985242  
ID AA1985242 standard; protein; 111 AA.

AC AA1985242;

XX 09-APR-1996 (first entry)

DT Humanised antibody variable light chain region (VZV HuV).

DE Antibody; varicella zoster virus; Ab; VSV; treatment; prevention;  
infection; detection; isolation; monoclonal antibody; MAb; mimetic;  
humanised; murine; heavy chain; light chain; variable region; CDR;  
NEW; REI.

OS Synthetic.

XX WO19531546-A1.

XX 23-NOV-1995.

XX 28-APR-1995; 95WO-US05305.

XX 28-APR-1994; 94US-0234133.

XX (SCOT-) SCOTGEN BIOPHARMACEUTICALS INC.

XX Carr FJ, Harris WJ, Moss MT, Wallace TP;

XX WPI; 1996-010932/01.

XX Varicella zoster virus gpIII antibody with humanised variable region  
PT - for treatment, prevention or diagnosis of varicella zoster virus  
PT infections

XX Claim 7; Page 31; 58pp; English.  
 XX Antibodies (Ab) comprising a humanised variable region specifically  
 CC binding to a varicella zoster virus (VZV) can be used to treat or  
 CC prevent VZV infection. They may optionally be attached to a  
 CC therapeutic agent. They may also be used, when labelled, to detect  
 CC VZV particles and infected cells in blood; for the removal/  
 CC neutralisation of infectious VZV in blood; for the selection/  
 CC isolation of human monoclonal Abs and for the design of (non-)  
 CC peptide mimetics with similar diagnostic and therapeutic uses. The  
 CC variable region CDR portion is derived from murine Ab 206 and the  
 CC heavy and light chain variable region frameworks are NEMM and REI  
 CC respectively. The framework may include at least one mutation that  
 CC improves binding specificity or affinity. The heavy and light chain  
 CC variable regions are preferably (VZVHu-VH, -VHKFIIS, -VHTLY,  
 CC -VHKVRSE or -VHS) and VZVHuVK respectively.  
 XX  
 SQ Sequence 111 AA;

Query Match 84.0%; Score 42; DB 17; Length 111;  
 Best Local Similarity 88.9%; Pred. No. 0.29;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPPT 9  
 III IIIII  
 Db 93 qhsrelpft 101

## RESULT 2

AAR85236  
 ID AAR85236 standard; Protein; 111 AA.

AC AAR85236;

DT 05-APR-1996 (first entry)

DE Murine 206 antibody variable light chain region.

XX Antibody; varicella zoster virus; Ab; VSV; treatment; prevention;  
 KW infection; detection; isolation; monoclonal antibody; MAB; mimetic;  
 KW humanised; murine; heavy chain; light chain; variable region; CDR;  
 KW NEMM; REI.

XX Mus musculus.

XX Key Location/Qualifiers  
 FH Binding-site 24..38  
 FT /label= CDR 1.  
 FT Binding-site 54..60  
 FT /label= CDR 2.  
 FT Binding-site 93..101  
 FT /label= CDR 3.

XX W09531546-A1.

XX 23-NOV-1995.

PF 28-APR-1995; 95WO-US05305.

PR 28-APR-1994; 94US-0234133.

XX (SCOT-) SCOTGEN BIOPHARMACEUTICALS INC.

XX Carr FJ, Harris WJ, Moss MT, Wallace TP;

XX WPI; 1996-010932/01.

XX N-PSDB; AAT06051.

XX Varicella zoster virus gpIII antibody with humanised variable region  
 PT - for treatment, prevention or diagnosis of varicella zoster virus  
 PT infections

XX Claim 2; Figure 2; 58pp; English.  
 XX Antibodies (Ab) comprising a humanised variable region specifically  
 CC binding to a varicella zoster virus (VZV) can be used to treat or  
 CC prevent VZV infection. They may optionally be attached to a  
 CC therapeutic agent. They may also be used, when labelled, to detect  
 CC VZV particles and infected cells in blood; for the removal/  
 CC neutralisation of infectious VZV in blood; for the selection/  
 CC isolation of human monoclonal Abs and for the design of (non-)  
 CC peptide mimetics with similar diagnostic and therapeutic uses. The  
 CC variable region CDR portion is derived from murine Ab 206 and the  
 CC heavy and light chain variable region frameworks are NEMM and REI  
 CC respectively. The framework may include at least one mutation that  
 CC improves binding specificity or affinity. The heavy and light chain  
 CC variable regions are preferably (VZVHu-VH, -VHKFIIS, -VHTLY,  
 CC -VHKVRSE or -VHS) and VZVHuVK respectively.  
 XX  
 SQ Sequence 111 AA;

Query Match 84.0%; Score 42; DB 17; Length 111;  
 Best Local Similarity 88.9%; Pred. No. 0.29;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPPT 9  
 III IIIII  
 Db 93 qhsrelpft 101

## RESULT 3

AAR75480  
 ID AAR75480 standard; Peptide; 9 AA.

AC AAR75480;

DT 07-FEB-1996 (first entry)

DE Mouse antibody light chain variable region CDR3 peptide #11.

XX Primer; amplification; PCR; mouse; kappa chain; heavy chain; Fab;  
 KW antibody; immunotolerance; animal; variegated display library; CDR;  
 KW variable region; antigen; immunorecessive; cell surface marker; foetal;  
 KW cancer; stem cell; variant; therapy; Alzheimer's disease; hybridoma;  
 KW familial hypercholesterolaemia; binding affinity;  
 KW complementarity determining region.

XX Synthetic.

XX W09515982-A2.

XX 15-JUN-1995.

PF 08-DEC-1994; 94WO-US14106.

XX 06-DEC-1994; 94US-0350400.

PR 08-DEC-1993; 93US-0164022.

XX (GENZ ) GENZYME CORP.

XX Barsomian G, Copeland DP, Hillhouse D, Johnson T;

XX WPI; 1995-224291/29.

XX Generating new antibodies specific for immunorecessive epitopes -  
 PT by selection from variegated V gene library cloned from  
 PT immuno:tolerance derived antibody repertoire, useful in diagnosis,  
 PT purificn. and therapy, e.g. of cancer

XX Claim 43; Page 92; 109pp; English.

XX Peptides AAR75462-92 are examples of complementarity determining regions  
 CC (CDRs) used to generate antibodies against immunorecessive antigens.

CC The CDRs are derived from the heavy and light chain variable regions of  
CC the antibodies FB3-2 (AAQ92500-1), F4-7 (AAQ92502-3) and H3-3  
CC (AAQ92504-5). The peptides AAR75470-81 represent CDR3 from the light  
CC chain variable regions. The CDR sequences were isolated from a  
CC variegated display library (VDL) of variable regions derived from a  
CC repertoire of antibodies from an immunotolerised animal. The VDL is  
CC generated by PCR amplifying the variable regions from the antibody coding  
CC sequences using the primers AAQ74153-74. The constructed antibodies are  
CC generated against an immunorecessive antigen e.g. a cell surface marker  
CC on a foetal, cancer or stem cell, which can differentiate between variant  
CC or related forms of the antigen. The antibodies generated can be used in  
CC the diagnosis, e.g. detection of the immunorecessive antigen, or in  
CC therapy e.g. of cancer, Alzheimer's disease or familial  
CC hypercholesterolaemia. The method of production of the antibody allows  
CC rapid and sensitive isolation of antibodies that would be difficult to  
CC isolate by standard methods. The antibodies produced have greater  
CC binding affinity than those produced by combinatorial/hybridoma methods.  
XX  
SQ Sequence 9 AA;

Query Match 80.0%; Score 40; DB 16; Length 9;  
Best Local Similarity 77.8%; Pred. No. 3.4e+05;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
DB 1 qhsweipft 9

RESULT 4  
AAR15440  
ID AAR15440 standard; Protein; 125 AA.

AC AAR15440;

DT 25-FEB-1992 (first entry)

DE Light chain variable region of MAB 2C2.

XX HRV; ICAM-1; antigen-binding fragment; inflammation;  
KW auto-immune disease.

XX Homo sapiens.

XX EP459577-A.

XX 04-DEC-1991.

XX 25-MAY-1991; 91EP-0201243.

XX 01-JUN-1990; 90US-0532001.

XX (MERI ) MERCK & CO INC.

XX Colonna RJ, Condra JH, Tomassini JE, Sardana VV;

XX WPI; 1991-355850/49.

XX Microbially expressed portions of monoclonal antibody - can block  
PT attachment of rhinovirus ligands to inter-cellular adhesion  
PT molecule (ICAM-1)

XX Claim 1; Page 20; 28pp; English.

XX This is one of six antibody fragments from MAB's specific for domain  
CC 1 of ICAM-1. The peptide fragments can be used to treat or prevent  
CC rhinovirus infection. See AAR15437-R15443.

XX Sequence 125 AA;

Query Match 80.0%; Score 40; DB 12; Length 125;

Best Local Similarity 77.8%; Pred. No. 0.85;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 QHSGELPFT 9  
DB 93 qhsweipft 101

RESULT 5

AAW03723

ID AAW03723 standard; Protein; 132 AA.

XX AAW03723;

DT 02-APR-1997 (first entry)

DE Anti-human gp39 MAB 39-1.7 light chain variable region.

XX Light chain; variable region; murine; mouse; anti-human; disease;  
KW glycoprotein 39; gp39; monoclonal; antibody; 39-1.7; hybridoma;  
KW diagnosis; inhibition; B-cell; activation; treatment; disorder;  
KW immune; autoimmune; allergic response; organ rejection; drug;  
KW graft versus host; cell imaging; tumour; targeted; delivery;  
KW targeted.

XX Mus musculus.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Peptide /label= sig\_peptide

FT Peptide 21..132

FT Region /label= mat\_peptide

FT Region 44..58

FT Region /note= "complementarity determining region 1"

FT Region /note= "complementarity determining region 2"

FT Region /note= "complementarity determining region 3"

XX WO9623071-A2.

XX 01-AUG-1996.

XX 26-JAN-1996; 96WO-US01119.

XX 26-JAN-1995; 95US-0379057.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Aruffo AA, Bajorath J, Gilliland LK, Gordon ML;

XX Harris LJ, Hollenbaugh D, Siadak AW;

XX WPI; 1996-362694/36.

XX N-PSDB; AAT35975.

XX Monoclonal antibodies specific for different epitope(s) on human  
PT gp39 - used for inhibiting B cell activation and for the diagnosis  
PT of various disorders, e.g. cancer, psoriasis etc..  
XX Claim 93; Fig 7A: 167pp; English.

CC The present sequence is the light chain variable region of the  
CC murine anti-human glycoprotein (gp) 39 monoclonal antibody (MAB)  
CC 39-1.7 (a member of the murine kappa II subfamily). The MAB was  
CC prep'd. by immunising a 6-8 week old BALB/c mouse with a gp39-CD8  
CC fusion protein, and 3 days later harvesting and fusing spleen and  
CC lymph cells to mouse melanoma cells, to produce an anti-human gp39  
CC MAB producing hybridoma. The MAB may be useful for diagnosing  
CC disease states, inhibiting B-cell activation and for treating  
CC immunological disorders, e.g. autoimmune disorders, allergic  
CC responses, organ rejection and graft versus host disease. It may  
CC also be used for imaging cells which express gp39 on their surface,  
CC e.g. tumour cells, and to target therapeutic agents to such cells.

CC The MAB inhibits the CD40/gp39 interaction, therefore limiting both  
 CC prim. and sec. responses to T-cell dependent antigens and Ab prodn.  
 CC specific to these antigens. A typical compsn. for intramuscular  
 CC injection pref. contains 50 mg of MAB in 1 ml of sterile buffered  
 CC water.

XX Sequence 132 AA;

Query Match 80.0%; Score 40; DB 17; Length 132;  
 Best Local Similarity 77.8%; Pred. No. 0.91;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9

Db 113 qhsweipft 121

RESULT 6

AAB42441

ID AAB42441 standard; Protein; 405 AA.

XX AC AAB42441;

XX 08-FEB-2001 (first entry)

XX Human ORFX ORF2205 polypeptide sequence SEQ ID NO:4410.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC76650.

XX Novel nucleic acids and peptides derived from open reading frame x,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 3603-3604; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;

CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
 CC antithyroid; and antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX protein in gene therapy  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus, lupus  
 CC hyperextension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX SQ Sequence 405 AA;

Query Match 78.0%; Score 39; DB 21; Length 405;

Best Local Similarity 87.5%; Pred. No. 5.1;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 8

Db 227 qhsgekpft 234

RESULT 7

AAR75475

ID AAR75475 standard; Peptide; 9 AA.

XX AC AAR75475;

XX 07-FEB-1996 (first entry)

XX Mouse antibody light chain variable region CDR3 peptide #6.

XX Primer; amplification; PCR; mouse; kappa chain; heavy chain; Fab;  
 KW antibody; immunotolerance; animal; variegated display library; CDR;  
 KW variable region; antigen; immunorecessive; cell surface marker; foetal;  
 KW cancer; stem cell; variant; therapy; Alzheimer's disease; hybridoma;  
 KW familial hypercholesterolaemia; binding affinity;  
 KW complementarity determining region.

XX Synthetic.

XX WO9515982-A2.

XX 15-JUN-1995.

XX 08-DEC-1994; 94WO-US14106.

XX 06-DEC-1994; 94US-0350400.

XX 08-DEC-1993; 93US-0164022.

XX (GENZ ) GENZYME CORP.

XX Barsomian G, Copeland DP, Hillhouse D, Johnson T;

XX WPI; 1995-224291/29.

XX Generating new antibodies specific for immunorecessive epitopes -  
 PT by selection from variegated V gene library cloned from  
 PT immunotolerance derived antibody repertoire, useful in diagnosis,  
 PT purifon. and therapy, e.g. of cancer

XX Claim 43; Page 92; 109pp; English.

XX Peptides AAR75462-92 are examples of complementarity determining regions  
 CC (CDRs) used to generate antibodies against immunorecessive antigens.

CC The CDRs are derived from the heavy and light chain variable regions of  
 CC the antibodies FB3-2 (AA092500-1), F4-7 (AA092502-3) and H3-3  
 CC (AA092504-5). The peptides AAR75470-81 represent CDR3 from the light  
 CC chain variable regions. The CDR sequences were isolated from a  
 CC variegated display library (VDL) of variable regions derived from a  
 CC repertoire of antibodies from an immunotolerised animal. The VDL is  
 CC generated by PCR amplifying the variable regions from the antibody coding  
 CC sequences using the primers AA074153-74. The constructed antibodies are  
 CC generated against an immunorecessive antigen e.g. a cell surface marker  
 CC on a foetal cancer or stem cell, which can differentiate between variant  
 CC or related forms of the antigen. The antibodies generated can be used in  
 CC the diagnosis, e.g. detection of the immunorecessive antigen, or in  
 CC therapy e.g. of cancer, Alzheimer's disease or familial  
 CC hypercholesterolaemia. The method of production of the antibody allows  
 CC rapid and sensitive isolation of antibodies that would be difficult to  
 CC isolate by standard methods. The antibodies produced have greater  
 CC binding affinity than those produced by combinatorial/hybridoma methods.

XX SQ Sequence 9 AA;

Query Match 74.0%; Score 37; DB 16; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 3.4e+05;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
 DB 1 qhsweipyt 9

RESULT 8  
 AAR85176  
 ID AAR85176 standard; peptide; 21 AA.

AC AAR85176;

DT 06-JUN-1996 (first entry)

DE GM-CSF receptor alpha-chain binding peptide.

XX Alpha-chain; GM-CSF; receptor; granulocyte; macrophage;  
 KW colony stimulating factor; binding assay; antagonist;  
 KW inflammation; autoimmune disease; agonist; upregulator;  
 KW immune system cell; CDR III; light chain; antibody 23.2.

XX Synthetic.

XX WO9529690-A1.

XX 09-NOV-1995.

XX 26-APR-1995; 95WO-US05160.

XX 29-APR-1994; 94US-0235404.

XX (UYPE-) UNIV PENNSYLVANIA.

XX (WIST-) WISTAR INST.

PI Kleber-Emmons T, Von Feldt JM, Weiner DB, Williams WV;

XX WPI; 1995-392919/50.

XX Identifying biologically active protein peptide mimics - by prepn.  
 PT of recombinant antibody libraries and synthesising complementary  
 PT determining region sequences.

XX Claim 13; Page 49; 70pp; English.

XX The present peptide binds the alpha-chain of the GM-CSF receptor,  
 CC and can therefore be used in binding assays in place of GM-CSF. It  
 CC can also be used as an antagonist, i.e. in the treatment of  
 CC inflammation or autoimmune diseases, or in the prodn. of GM-CSF  
 CC agonists, i.e. upregulators for the prodn. of immune system cells.

CC The peptide is derived from CDR III of the recombinant antibody  
 CC light chain analogue of GM-CSF, 23.2.

XX SQ Sequence 21 AA;

Query Match 74.0%; Score 37; DB 16; Length 21;  
 Best Local Similarity 77.8%; Pred. No. 0.5;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
 DB 2 qhsrelpwt 10

RESULT 9

AAR25721  
 ID AAR25721 standard; Protein; 111 AA.

AC AAR25721;

DT 13-JAN-1993 (first entry)

DE Humanised VL region of the mouse Fd79 antibody.

XX Murine; immunoglobulin; CDR; HSV; non immunogenic; herpes simplex  
 KW virus; HSV; light chain; variable region; framework; human; Pom.

XX Mus musculus.

XX Key Location/Qualifiers

FT Region 24..38

FT /note= "CDR"

FT Region 54..60

FT /note= "CDR"

FT Region 93..100

FT /note= "CDR"

FT Misc-difference 9

FT /note= "mutated residue"

FT Misc-difference 45

FT /note= "mutated residue"

FT Misc-difference 46

FT /note= "mutated residue"

FT Misc-difference 53

FT /note= "mutated residue"

FT Misc-difference 81

FT /note= "mutated residue"

FT Misc-difference 83

FT /note= "mutated residue"

XX WO9211018-A.

XX 09-JUL-1992.

XX 19-DEC-1991; 91WO-US09711.

XX 19-DEC-1990; 90US-0634278.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Co MS, Coeligh KL, Landolfi NF, Queen CL, Schneider WP;

XX WPI; 1992-249842/30.

XX New immunoglobulin(s) having murine CDRs in human framework  
 PT regions - have lower antigenicity; useful for treating e.g. HSV,  
 PT CMV, T-cell disorders, myeloid disorders and auto-immune  
 PT conditions

XX Claim 19; Fig 143; 141pp; English.

XX The sequence shows the humanised mature light chain variable  
 CC region of the mouse Fd79 antibody. Murine CDRs were used in a

CC human Pom framework to produce a pure humanised immunoglobulin (Ig)  
 CC which is capable of binding to a herpes simplex virus-specific  
 CC epitope. The Ig is non immunogenic, due to the human framework,  
 CC and has a strong affinity for its predetermined antigen. They can  
 CC be produced in large quantities via recombinant DNA and monoclonal  
 CC antibody technology. The humanised Igs may be used alone or in  
 CC combination with chemotherapeutic agents such as non-steroidal  
 CC anti-inflammatory drugs or immunosuppressants.  
 CC See also AAR25722-32.  
 XX  
 XX

SQ Sequence 111 AA;

Query Match 74.0%; Score 37; DB 13; Length 111;  
 Best Local Similarity 66.7%; Pred. No. 3.1;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QHSGELPFT 9  
 ||| |::|  
 Db 93 qhsweipyt 101

RESULT 10  
 AAB69661  
 ID AAB69661 standard; Protein; 111 AA.

AC AAB69661;

XX 30-APR-2001 (first entry)

XX Murine Fd79 antibody light chain SEQ ID NO: 46.

XX Humanised immunoglobulin; mouse; human; antibody; heavy chain; diabetes;  
 KW light chain; graft versus host disease; transplant; autoimmune disease;  
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW myasthenia gravis; herpes infection; myeloid leukaemia; CMV infection.

XX Mus sp.

XX US6180370-B1.

XX 30-JAN-2001.

XX 07-JUN-1995; 95US-0484537.

XX 28-DEC-1988; 88US-0290975.

PR 13-FEB-1989; 89US-0310252.

PR 28-SEP-1990; 90US-0590274.

PR 19-DEC-1990; 90US-0634278.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Queen CL, Selick HE;

XX WPI; 2001-190856/19.

XX Producing humanized immunoglobulin, involves producing a cell  
 PT containing DNA segments encoding humanized heavy and light chain  
 PT variable regions, and expressing the DNA segments in the cell -

XX Example 6; Fig 2; 145pp; English.

XX The present invention describes a method of producing humanised  
 CC immunoglobulins involving expressing in a cell a nucleic acid encoding a  
 CC humanised version of an immunoglobulin. This is obtained by comparing a  
 CC donor and human immunoglobulin and producing a combined antibody which  
 CC contains part of each. These are useful in the treatment of  
 CC graft-versus-host disease, transplant rejection, autoimmune diseases such  
 CC as diabetes, rheumatoid arthritis, myasthenia gravis, multiple sclerosis  
 CC and systemic lupus erythematosus, herpes infections, CMV virus infections  
 CC and myeloid leukaemia. The present sequence is an antibody used to  
 CC demonstrate the method of the invention.

SQ Sequence 111 AA;

Query Match 74.0%; Score 37; DB 22; Length 111;  
 Best Local Similarity 66.7%; Pred. No. 3.1;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QHSGELPFT 9  
 ||| |::|  
 Db 93 qhsweipyt 101

RESULT 11

AAB69662

ID AAB69662 standard; Protein; 111 AA.

XX AAB69662;

XX 30-APR-2001 (first entry)

XX Humanised Fd79 antibody light chain SEQ ID NO: 47.

XX Humanised immunoglobulin; mouse; human; antibody; heavy chain; diabetes;  
 KW light chain; graft versus host disease; transplant; autoimmune disease;  
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW myasthenia gravis; herpes infection; myeloid leukaemia; CMV infection.

XX Mus sp.

XX Homo sapiens.

XX US6180370-B1.

XX 30-JAN-2001.

XX 07-JUN-1995; 95US-0484537.

XX 28-DEC-1988; 88US-0290975.

PR 13-FEB-1989; 89US-0310252.

PR 28-SEP-1990; 90US-0590274.

PR 19-DEC-1990; 90US-0634278.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Queen CL, Selick HE;

XX WPI; 2001-190856/19.

XX Producing humanized immunoglobulin, involves producing a cell  
 PT containing DNA segments encoding humanized heavy and light chain  
 PT variable regions, and expressing the DNA segments in the cell -

XX Example 6; Fig 2; 145pp; English.

XX The present invention describes a method of producing humanised  
 CC immunoglobulins involving expressing in a cell a nucleic acid encoding a  
 CC humanised version of an immunoglobulin. This is obtained by comparing a  
 CC donor and human immunoglobulin and producing a combined antibody which  
 CC contains part of each. These are useful in the treatment of  
 CC graft-versus-host disease, transplant rejection, autoimmune diseases such  
 CC as diabetes, rheumatoid arthritis, myasthenia gravis, multiple sclerosis  
 CC and systemic lupus erythematosus, herpes infections, CMV virus infections  
 CC and myeloid leukaemia. The present sequence is an antibody used to  
 CC demonstrate the method of the invention.

XX Sequence 111 AA;

Query Match 74.0%; Score 37; DB 22; Length 111;  
 Best Local Similarity 66.7%; Pred. No. 3.1;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QHSGELPFT 9  
 ||| |::|



Db 93 qhsweipyt 101

RESULT 12

AAR75457  
ID AAR75457 standard; Protein; 218 AA.

AC AAR75457;  
XX

DT 07-FEB-1996 (first entry)  
XX

DE Mouse antibody FB3-2 light chain variable region protein sequence.  
XX

XX Primer; amplification; PCR; mouse; kappa chain; heavy chain; Fab;  
KW antibody; immunotolerance; animal; variegated display library;  
KW variable region; antigen; immunorecessive; cell surface marker; foetal;  
KW cancer; stem cell; variant; therapy; Alzheimer's disease; hybridoma;  
KW familial hypercholesterolaemia; binding affinity.  
XX

OS Mus musculus.  
XX

PN WO9515982-A2.  
XX

PD 15-JUN-1995.  
XX

PF 08-DEC-1994; 94WO-US14106.  
XX

PR 06-DEC-1994; 94US-0350400.  
XX

PR 08-DEC-1993; 93US-0164022.  
XX

PA (GENZ ) GENZYME CORP.  
XX

PI Barsomian G, Copeland DP, Hillhouse D, Johnson T;  
XX

DR WPI; 1995-224291/29.  
XX

DR N-PSDB; AAQ92501.  
XX

XX Generating new antibodies specific for immunorecessive epitopes  
PT by selection from variegated V gene library cloned from  
PT immunotolerance derived antibody repertoire, useful in diagnosis,  
PT purificn. and therapy, e.g. of cancer  
XX

PS Disclosure; Page 76-77; 109pp; English.  
XX

XX The sequence of the light chain variable region from the mouse antibody  
CC FB3-2. This sequence was isolated from a variegated display library  
CC (VDL) of variable regions derived from a repertoire of antibodies from an  
CC immunotolerant animal. The VDL is generated by PCR amplifying the  
CC variable regions from the antibody coding sequences using the primers  
CC AAQ74153-74. The variable regions, esp the complementarity determining  
CC regions (CDR; see AAR75462-93 for examples of CDRs) from the  
CC immunotolerant animals' antibodies are used to construct an antibody  
CC against a immunorecessive antigen e.g. a cell surface marker on a foetal,  
CC cancer or stem cell, which can differentiate between variant or related  
CC forms of the antigen. The antibodies generated can be used in the  
CC diagnosis, e.g. detection of the immunorecessive antigen, or in therapy  
CC e.g. of cancer, Alzheimer's disease or familial hypercholesterolaemia.  
CC The method of production of the antibody allows rapid and sensitive  
CC isolation of antibodies that would be difficult to isolate by standard  
CC methods. The antibodies produced have greater binding affinity than  
CC those produced by combinatorial/hybridoma methods.  
XX

SQ Sequence 218 AA;

Query Match 74.0%; Score 37; DB 16; Length 218;  
Best Local Similarity 66.7%; Pred. No. 6.6;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
||| |::|  
Db 93 qhsweipyt 101

RESULT 13

AAR75459

XX AAR75459 standard; Protein; 218 AA.

AC AAR75459;  
XX

DT 07-FEB-1996 (first entry)  
XX

DE Mouse antibody F4-7 light chain variable region protein sequence.  
XX

XX Primer; amplification; PCR; mouse; kappa chain; heavy chain; Fab;  
KW antibody; immunotolerance; animal; variegated display library;  
KW variable region; antigen; immunorecessive; cell surface marker; foetal;  
KW cancer; stem cell; variant; therapy; Alzheimer's disease; hybridoma;  
KW familial hypercholesterolaemia; binding affinity.  
XX

OS Mus musculus.  
XX

PN WO9515982-A2.  
XX

PD 15-JUN-1995.  
XX

PF 08-DEC-1994; 94WO-US14106.  
XX

PR 06-DEC-1994; 94US-0350400.  
XX

PR 08-DEC-1993; 93US-0164022.  
XX

PA (GENZ ) GENZYME CORP.  
XX

PI Barsomian G, Copeland DP, Hillhouse D, Johnson T;  
XX

DR WPI; 1995-224291/29.  
XX

DR N-PSDB; AAQ92503.  
XX

XX Generating new antibodies specific for immunorecessive epitopes  
PT by selection from variegated V gene library cloned from  
PT immunotolerance derived antibody repertoire, useful in diagnosis,  
PT purificn. and therapy, e.g. of cancer  
XX

PS Disclosure; Page 80-81; 109pp; English.  
XX

XX The sequence of the light chain variable region from the mouse antibody  
CC F4-7. This sequence was isolated from a variegated display library (VDL)  
CC of variable regions derived from a repertoire of antibodies from an  
CC immunotolerant animal. The VDL is generated by PCR amplifying the  
CC variable regions from the antibody coding sequences using the primers  
CC AAQ74153-74. The variable regions, esp the complementarity determining  
CC regions (CDR; see AAR75462-93 for examples of CDRs) from the  
CC immunotolerant animals' antibodies are used to construct an antibody  
CC against a immunorecessive antigen e.g. a cell surface marker on a foetal,  
CC cancer or stem cell, which can differentiate between variant or related  
CC forms of the antigen. The antibodies generated can be used in the  
CC diagnosis, e.g. detection of the immunorecessive antigen, or in therapy  
CC e.g. of cancer, Alzheimer's disease or familial hypercholesterolaemia.  
CC The method of production of the antibody allows rapid and sensitive  
CC isolation of antibodies that would be difficult to isolate by standard  
CC methods. The antibodies produced have greater binding affinity than  
CC those produced by combinatorial/hybridoma methods.  
XX

SQ Sequence 218 AA;

Query Match 74.0%; Score 37; DB 16; Length 218;  
Best Local Similarity 66.7%; Pred. No. 6.6;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
||| |::|  
Db 93 qhsweipyt 101

RESULT 14

AAB10000  
ID AAB10000 standard; Protein; 9 AA.

XX AAB10000;  
XX 01-NOV-2000 (first entry)  
XX H. pylori HSP60-binding antibody light chain CDR3 protein fragment.  
XX Acid-resistant microorganism; detection; faecal; intestine; infection;  
KW monoclonal antibody; heat shock protein; HSP60; light chain; CDR;  
KW complementarity determining region.

XX Unidentified.

XX WO200026671-A1.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-EP08212.

XX 29-OCT-1998; 98EP-0120517.

PR 06-NOV-1998; 98EP-0120687.

XX (CONN-) CONNEX GMBH.

XX Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;

PI Ringeis A;

XX WPI; 2000-365747/31.

DR N-PSDB; AAA40156.

XX Detecting infection by acid-fast microbes for diagnosis of Helicobacter  
PT Pylori, comprises reacting a faecal sample with two binding reagents for  
PT antigens that survive intestinal passage -

XX Claim 20; Page 20; 84pp; German.

XX This invention describes a novel method for the detection of a mammalian  
CC infection by an acid-resistant microorganism (A) by treating a faecal  
CC sample with at least two different monoclonal antibodies (MAB) (or their  
CC fragments or derivatives) or aptamers (collectively (I)) and detecting  
CC formation of a complex (C) between (I) and the corresponding antigen of  
CC (A). The first and second (I) bind to epitopes of different antigens  
CC (Ag). These epitopes are present, after passage through the intestines,  
CC in at least some mammals, and have either: (i) their native structure,  
CC or (ii) a structure against which an antibody is produced by an animal  
CC infected or immunized with (A), or its extract, lysate, derived protein  
CC or fragment, or with a synthetic peptide. Practically all mammals display  
CC at least one of the specified epitopes. The method is used to detect  
CC infection by acid-fast bacteria, particularly of the genera Helicobacter,  
CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,  
CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used  
CC therapeutically. The method is direct and non-invasive, and provides an  
CC inexpensive and easily standardizable diagnosis, despite possible  
CC degradation of antigens during passage through the intestines. This  
CC sequence represents a fragment of a H. pylori heat shock protein,  
CC HSP60-binding antibody light chain complementarity determining region  
CC CDR3 which is used to illustrate the method of the invention.

XX Sequence 9 AA;

Query Match 72.0%; Score 36; DB 21; Length 9;  
Best Local Similarity 77.8%; Pred. No. 3.4e+05;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9

Db 1 qhsrelplt 9

RESULT 15

AAR60528  
ID AAR60528 standard; Protein; 111 AA.

XX AAR60528;

XX 07-APR-1995 (first entry)

XX Light chain variable region of monoclonal antibody 5E8.

XX Secretion sequence; vector; antibody; single chain antibody;  
KW promoter; Bacillus subtilis; ss.

XX Synthetic.

XX EP614982-A.

XX 14-SEP-1994.

XX 05-MAR-1994; 94EP-0200564.

PR 10-MAR-1993; 93IT-OMI0456.

XX (ENIE) ENRICERCH SPA.

XX Cuzzoni A, De FERRA F, Grandi G, Tortora O, Tosi C;

XX WPI; 1994-281209/35.

DR N-PSDB; AAQ71286.

XX Vector for secretion of single chain antibody from Bacillus  
PT subtilis - contg. neutral protease promoter and specific leader  
PT sequence, providing high yield of soluble antibody for diagnostic  
PT or therapeutic use

XX Example 3; Figure 2; 27pp; English.

XX The sequence encoding the variable region of the light chain of the  
CC monoclonal antibody 5E8 (specific for the alpha subunit of human  
CC gonadotrophin) was amplified for its use in a recombinant vector  
CC which also comprised (1) the promoter of the gene for the neutral  
CC protease of Bacillus subtilis BGSCL341, (2) the secretion sequence  
CC described in AAQ71278. The amplified fragment was used to produce a  
CC DNA sequence encoding an antibody in single molecule form with the  
CC sequence VH/VK-L-VK/VH-(TAG)n, where VH and VK are the variable  
CC regions of the heavy and light chains of the antibody and L is the  
CC linker of sequence Val-Ser-Ser-(Gly(4)-Ser)<sub>3</sub>. TAG is a peptide  
CC recognised by polyclonal antibodies directed towards the same peptide,  
CC n is 1 or 0. The vector is used for the production of single chain  
CC antibodies which have improved pharmacokinetic properties and can be  
CC produced more economically than monoclonal antibodies. The vector  
CC allows production of antibodies in completely soluble form with  
CC secretion in high yield.

XX Sequence 111 AA;

Query Match 72.0%; Score 36; DB 15; Length 111;  
Best Local Similarity 77.8%; Pred. No. 5.1;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9

Db 93 qhsrelplt 101

Search completed: June 28, 2001, 16:14:39  
Job time: 1327 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:01:15 ; Search time 138.34 Seconds  
(without alignments)  
1.311 Million cell updates/sec

Title: US-09-724-406-32  
Perfect score: 50  
Sequence: 1 QHSGELPPT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 2014635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	80.0	9	PCT-US94-14106-46	Sequence 46, Appl
2	40	80.0	132	US-08-379-057-16	Sequence 16, Appl
3	37	74.0	9	PCT-US94-14106-40	Sequence 40, Appl
4	37	74.0	21	US-08-712-212-2	Sequence 2, Appl
5	37	74.0	21	PCT-US95-05160-2	Sequence 2, Appl
6	37	74.0	111	US-07-634-278-46	Sequence 46, Appl
7	37	74.0	111	US-07-634-278-47	Sequence 47, Appl
8	37	74.0	111	US-08-477-728-46	Sequence 46, Appl
9	37	74.0	111	US-08-477-728-47	Sequence 47, Appl
10	37	74.0	111	US-08-474-040-46	Sequence 46, Appl
11	37	74.0	111	US-08-474-040-47	Sequence 47, Appl
12	37	74.0	111	US-08-487-200-46	Sequence 46, Appl
13	37	74.0	111	US-08-487-200-47	Sequence 47, Appl
14	37	74.0	111	US-08-484-537-46	Sequence 46, Appl
15	37	74.0	111	US-08-484-537-47	Sequence 47, Appl
16	37	74.0	111	PCT-US94-14106-53	Sequence 53, Appl
17	37	74.0	218	PCT-US94-14106-57	Sequence 57, Appl
18	37	74.0	453	5206152-7	Patent No. 5206152
19	36	72.0	77	US-08-570-227A-6	Sequence 6, Appl
20	36	72.0	77	US-08-077-991-8	Sequence 8, Appl
21	36	72.0	111	US-08-207-169A-4	Sequence 4, Appl
22	36	72.0	640	US-09-262-773-4	Sequence 4, Appl
23	36	72.0	648	US-09-262-773-2	Sequence 2, Appl
24	34	68.0	56	US-08-711-417C-186	Sequence 186, App
25	34	68.0	56	US-08-711-417C-187	Sequence 187, App
26	34	68.0	56	US-08-711-417C-188	Sequence 188, App
27	34	68.0	56	US-08-711-417C-189	Sequence 189, App

28	34	68.0	77	2	US-08-570-227A-4	Sequence 4, Appl
29	34	68.0	77	4	US-09-077-991-6	Sequence 6, Appl
30	34	68.0	84	2	US-08-459-568-62	Sequence 62, Appl
31	34	68.0	84	2	US-08-399-411-62	Sequence 62, Appl
32	34	68.0	84	3	US-08-516-859A-62	Sequence 62, Appl
33	34	68.0	93	1	US-08-083-947-12	Sequence 12, Appl
34	34	68.0	93	5	PCT-US94-07088-12	Sequence 12, Appl
35	34	68.0	334	4	US-08-711-417C-201	Sequence 201, App
36	34	68.0	338	2	US-08-933-750C-4	Sequence 4, Appl
37	34	68.0	338	4	US-09-234-613-4	Sequence 4, Appl
38	34	68.0	390	4	US-08-711-417C-199	Sequence 199, App
39	34	68.0	431	4	US-08-711-417C-195	Sequence 195, App
40	34	68.0	432	4	US-08-711-417C-197	Sequence 197, App
41	34	68.0	461	4	US-08-711-417C-196	Sequence 196, App
42	34	68.0	470	2	US-08-465-590-153	Sequence 153, App
43	34	68.0	470	4	US-08-283-300A-13	Sequence 13, App
44	34	68.0	470	4	US-08-711-417C-153	Sequence 153, App
45	34	68.0	470	5	PCT-US95-09345-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
PCT-US94-14106-46  
; Sequence 46, Application PC/TUS9414106  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Process for Generating Specific Antibodies  
; NUMBER OF SEQUENCES: 61  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/14106  
; FILING DATE:  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
PCT-US94-14106-46

Query Match 80.0%; Score 40; DB 5; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.4e+05;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPPT 9  
||| |  
Db 1 QHSWEIPPT 9

RESULT 2  
US-08-379-057-16  
; Sequence 16, Application US/08379057  
; Patent No. 5876950  
; GENERAL INFORMATION:  
; APPLICANT: Siadak, Anthony W.  
; APPLICANT: Hollibaugh, Diane L.  
; APPLICANT: Gilliland, Lisa K.  
; APPLICANT: Gordon, Marcia L.  
; APPLICANT: Bajorath, Jurgen  
; APPLICANT: Aruffo, Alejandro A.  
; TITLE OF INVENTION: Monoclonal Antibodies Specific For  
; DIFFERENT EPITOPES OF HUMAN gp39 AND METHODS FOR THEIR USE  
; TITLE OF INVENTION: In Diagnosis and Therapy  
; NUMBER OF SEQUENCES: 57

1 OHSGELPFT 9 Ov

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: Norris  
STREET: One Liberty Place, 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05160  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/235,404  
FILING DATE: 29-APR-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: UPN-2245  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-05160-2

Query Match 74.0%; Score 37; DB 5; Length 21;  
Best Local Similarity 77.8%; Pred. No. 0.24;  
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 1 QHSGLPFT 9  
Db 2 QHSRELPWT 10

RESULT 6  
US-07-634-278-46  
; Sequence 46, Application US/07634278  
; Patent No. 5530101  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO, Man Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLFI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/634,278  
; FILING DATE: 19-DEC-1990

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-634-278-46

Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGLPFT 9  
Db 93 QHSWEIPYT 101

RESULT 7  
US-07-634-278-47  
; Sequence 47, Application US/07634278  
; Patent No. 5530101  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO, Man Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLFI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/634,278  
; FILING DATE: 19-DEC-1990  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252  
; FILING DATE: 13-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/290,975

; FILING DATE: 28-DEC-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-002600  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-634-278-47

Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPFT 9  
Db 93 QHSWEIPYT 101

RESULT 8  
US-08-477-728-46  
; Sequence 46, Application US/08477728  
; Patent No. 5585089  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: SCHNEIDER, William P.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,728  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/634,278  
; FILING DATE: 19-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252  
; FILING DATE: 13-FEB-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-002600  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 46:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-477-728-46

Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPFT 9  
Db 93 QHSWEIPYT 101

RESULT 9  
US-08-477-728-47  
; Sequence 47, Application US/08477728  
; Patent No. 5585089  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,728  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/634,278  
; FILING DATE: 19-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252  
; FILING DATE: 13-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/290,975  
; FILING DATE: 28-DEC-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-002600  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-477-728-47



Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
||| |::|  
Db 93 QHSWEIPYT 101

RESULT 10  
US-08-474-040-46  
; Sequence 46, Application US/08474040  
; Patent No. 5693761  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO, Man Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLFI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,040  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/634,278  
; FILING DATE: 19-DEC-1990  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252  
; FILING DATE: 13-FEB-1989  
; APPLICATION NUMBER: US 07/290,975  
; FILING DATE: 28-DEC-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-002600  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-474-040-46

Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
||| |::|  
Db 93 QHSWEIPYT 101

RESULT 11  
US-08-474-040-47  
; Sequence 47, Application US/08474040  
; Patent No. 5693761  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO, Man Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLFI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,040  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/634,278  
; FILING DATE: 19-DEC-1990  
; APPLICATION NUMBER: US 07/590,274  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/310,252  
; FILING DATE: 13-FEB-1989  
; APPLICATION NUMBER: US 07/290,975  
; FILING DATE: 28-DEC-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-002600  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 111 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-474-040-47

Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
||| |::|  
Db 93 QHSWEIPYT 101

RESULT 12  
US-08-487-200-46  
; Sequence 46, Application US/08487200  
; Patent No. 5693762  
; GENERAL INFORMATION:

APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELINGH, Kathleen L.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,200  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002610  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-487-200-46

Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPFT 9  
Db 93 QHSWEIPT 101

RESULT 13  
US-08-487-200-47  
Sequence 47, Application US/08487200  
Patent No. 5693762  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELINGH, Kathleen L.  
APPLICANT: SELICK, Harold E.

TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,200  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002610  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-487-200-47

Query Match 74.0%; Score 37; DB 1; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPFT 9  
Db 93 QHSWEIPT 101

RESULT 14  
US-08-484-537-46  
Sequence 46, Application US/08484537  
Patent No. 6180370  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELINGH, Kathleen L.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Khourie and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto

STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,537  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/634,278  
FILING DATE: 19-DEC-1990  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-484-537-46

Query Match 74.0%; Score 37; DB 4; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
Db 93 QHSWEIPYT 101

RESULT 15  
US-08-484-537-47  
Sequence 47, Application US/08484537  
Patent No. 6180370  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO. Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELINGH, Kathleen L.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,537  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/634,278  
FILING DATE: 19-DEC-1990  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-484-537-47

Query Match 74.0%; Score 37; DB 4; Length 111;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
Db 93 QHSWEIPYT 101

Search completed: June 28, 2001, 16:01:15  
Job time: 523 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:58:46 ; Search time 234.85 seconds  
(without alignments)  
2.919 Million cell updates/sec

Title: US-09-724-406-32  
Perfect score: 50  
Sequence: 1 QHSGELPFT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	80.0	102	PH1077	Ig light chain V r
2	39	78.0	223	S06576	finger protein (cl
3	39	78.0	405	T42663	hypothetical prote
4	39	78.0	594	T16387	hypothetical prote
5	37	74.0	101	S59640	Ig light chain V r
6	37	74.0	111	S09963	Ig kappa chain V-J
7	37	74.0	210	A56169	Ig kappa chain V r
8	37	74.0	384	I40867	hypothetical prote
9	37	74.0	1143	B84431	probable receptor
10	36	72.0	52	F43284	finger protein ZNF
11	36	72.0	111	KVMS75	Ig kappa chain V r
12	36	72.0	111	KVMS85	Ig kappa chain V r
13	36	72.0	112	S06562	finger protein (cl
14	36	72.0	138	S06571	finger protein (cl
15	36	72.0	140	S08574	finger protein (cl
16	36	72.0	157	S03632	finger protein (cl
17	36	72.0	169	S06580	finger protein (cl
18	36	72.0	196	S06558	finger protein (cl
19	36	72.0	196	S06581	finger protein (cl
20	36	72.0	196	S06561	finger protein (cl
21	36	72.0	196	S06555	finger protein (cl
22	36	72.0	214	S00832	finger protein (cl
23	36	72.0	218	S68241	Ig kappa chain V r
24	36	72.0	218	JC5810	monoclonal antibody
25	36	72.0	247	S06553	finger protein (cl
26	36	72.0	252	S06567	finger protein (cl
27	36	72.0	280	S06572	finger protein (cl
28	36	72.0	336	S06573	finger protein (cl
29	36	72.0	336	S06578	finger protein (cl

30 36 72.C 337 2 S06566 finger protein (cl  
31 36 72.C 349 1 RGECCGL nitrogen regulatio  
32 36 72.C 349 2 A24114 nitrogen regulatio  
33 36 72.C 349 2 F86074 hypothetical prote  
34 36 72.C 353 2 S65086 finger protein XFO  
35 36 72.C 420 2 S65084 finger protein XFG  
36 36 72.C 439 2 S06556 finger protein (cl  
37 36 72.C 445 2 S32036 finger protein XFG  
38 36 72.C 462 2 I51699 gene XGF 5.1C prot  
39 36 72.C 474 2 I54338 zinc finger protei  
40 36 72.C 484 2 I39200 zinc finger protei  
41 36 72.C 493 2 S32037 finger protein XFG  
42 36 72.C 591 2 S65088 finger protein XFO  
43 36 72.C 634 2 T47156 hypothetical prote  
44 36 72.C 675 2 S51037 zinc-finger protei  
45 35 70.0 131 2 PL0207 anti-idiotypic ant

#### ALIGNMENTS

RESULT 1

PH1077

Ig light chain V region (clone 17s-c4) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000

C:Accession: PH1077

R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective

A:Reference number: PH0971; MUID:92381444

A:Accession: PH1077

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-102 <TIL>

A:Experimental source: B cell, strain [N2B x NZW]F1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: immunoglobulin

F:16-94/Domain: immunoglobulin homology <IMM>

Query Match 80.0%; Score 40; DB 2; Length 102;  
Best Local Similarity 77.8%; Pred. No. 0.62;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9

DB 93 QHSWEIPT 101

RESULT 2

S06576

finger protein (clone XlcGF52-1) - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 31-Dec-1993

C:Accession: S06576

R:Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeting, A.; Kno

J. Mol. Biol. 208, 639-659, 1989

A:Title: Second-order repeats in Xenopus laevis finger proteinins.

A:Reference number: S05632; MUID:90040698

A:Accession: S06576

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-223 <NIE>

C:Keywords: DNA binding; zinc finger

Query Match 78.0%; Score 39; DB 2; Length 223;  
Best Local Similarity 87.5%; Pred. No. 2.3;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9

|||||

Db 195 HSGEKPFT 202

RESULT 3

T42663

hypothetical protein DKFZp434N0615.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000

C:Accession: T42663

R:Boecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, November 1999

A:Reference number: Z22230

A:Accession: T42663

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-405 <HAA>

A:Cross-references: EMBL:AL133062

A:Experimental source: adult testis; clone DKFZp434N0615

C:Genetics:

A:Note: DKFZp434N0615.1

Query Match

Best Local Similarity 78.0%; Score 39; DB 2; Length 405;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 8

Db 227 QHSGEKPF 234

RESULT 4

T16387

hypothetical protein F47E1.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16387

R:Favell, T.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid F47E1.

A:Reference number: Z18505

A:Accession: T16387

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-594 <FAV>

A:Cross-references: EMBL:U39993; NID:g1055098; PID:g1055099; PIDN:AAA81086.1; CESP:F47E1

C:Genetics:

A:Gene: CESP:F47E1.3

A:Introns: 20/1; 54/3; 78/3; 104/3; 163/3; 186/3; 297/2; 352/1; 396/3; 457/1; 502/2; 541

Query Match

Best Local Similarity 78.0%; Score 39; DB 2; Length 594;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 8

Db 573 QHSGEKPF 580

RESULT 5

S59640

Ig light chain V region N10 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 21-Jan-2000

C:Accession: S59640

R:Bosart-Whitaker, P.; Chang, C.Y.Y.; Novotny, J.; Benjamin, D.C.; Sheriff, S.

J. Mol. Biol. 253, 559-575, 1995

A:Title: The crystal structure of the antibody N10-staphylococcal nuclease complex at 2.

A:Reference number: S59639; MUID:96068846

A:Accession: S59640

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-101 <BOS>

A:Cross-references: EMBL:U25122

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: immunoglobulin

F:6-84/Domain: immunoglobulin homology <IMM>

Query Match 74.0%; Score 37; DB 2; Length 101;

Best Local Similarity 66.7%; Pred. No. 2.5;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9

Db 83 QHSEIPT 91

RESULT 6

S09963

Ig kappa chain V-J region (103-7E) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 21-Jan-2000

C:Accession: S09963

R:Reininger, L.; Shibata, T.; Ozaki, S.; Shirai, T.; Jaton, J.C.; Izui, S.

Eur. J. Immunol. 20, 771-777, 1990

A:Title: Variable region sequences of pathogenic anti-mouse red blood cell autoantibo

A:Reference number: S09955; MUID:90269328

A:Accession: S09963

A:Molecule type: mRNA

A:Residues: 1-111 <REI>

A:Cross-references: EMBL:X51851

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:16-94/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 74.0%; Score 37; DB 2; Length 111;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9

Db 93 QHSEIPT 101

RESULT 7

A56169

Ig kappa chain V region (clone 23.2) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 11-Jan-2000

C:Accession: A56169

R:Monfardini, C.; Kieber-Emmons, T.; VonFeldt, J.M.; O'Malley, B.; Rosenbaum, H.; God

J. Biol. Chem. 270, 6628-6638, 1995

A:Title: Recombinant antibodies in bioactive peptide design.

A:Reference number: A56169; MUID:95204454

A:Accession: A56169

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-210 <MON>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match

Best Local Similarity 74.0%; Score 37; DB 2; Length 210;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9

Db 93 QHSEIPT 101

RESULT 8

I40867

hypothetical protein 2 - Clostridium perfringens  
 C:Species: Clostridium perfringens  
 C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Oct-1999  
 C:Accession: I40867  
 R:Traving, C.; Schauer, R.; Roggentin, P.  
 Glycocon]. J. 11, 141-151, 1994  
 A:Title: Gene structure of the 'large' sialidase isoenzyme from Clostridium perfringens  
 A:Reference number: I40865; MUID:95102306  
 A:Accession: I40867  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-384 <RES>  
 A:Cross-references: EMBL:X87369; NID:g853805; PIDN:CAA60797.1; PID:g853808

Query Match 74.0%; Score 37; DB 2; Length 384;  
 Best Local Similarity 66.7%; Pred. No. 11;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
 : I I I I I I  
 Db 157 KHIGDLPT 165

RESULT 9  
 B84431  
 Probable receptor protein kinase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: B84431  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: B84420; MUID:20083487  
 A:Accession: B84431  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1143 <STO>  
 A:Cross-references: GB:AE002093; NID:g4406778; PIDN:AAD20088.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g01950  
 A:Map position: 2

Query Match 74.0%; Score 37; DB 2; Length 1143;  
 Best Local Similarity 77.8%; Pred. No. 35;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
 : I I I I I I  
 Db 646 QLSGEIPT 654

RESULT 10  
 F43284  
 finger protein ZNF52 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
 C:Accession: F43284  
 R:Lichter, P.; Bray, P.; Ried, T.; Dawid, I.B.; Ward, D.C.  
 Genomics 13, 999-1007, 1992  
 A:Title: Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile si  
 A:Reference number: A43284; MUID:92372070  
 A:Accession: F43284  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-52 <LIC>  
 A:Cross-references: GB:N88361; NID:g340459; PID:g553813  
 A:Note: sequence extracted from NCB1 backbone (NCBIN:111644, NCBIP:111646)  
 C:Keywords: zinc finger

Query Match 72.0%; Score 36; DB 2; Length 52;  
 Best Local Similarity 75.0%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
 : I I I I I I  
 Db 21 HSGEKPYT 28

RESULT 11  
 KVMS75  
 Ig kappa chain V region (PC7175) - mouse (tentative sequence)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 31-Mar-2000  
 C:Accession: B01938; A01938  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: B01938  
 A:Molecule type: protein  
 A:Residues: 1-111 <WEI>  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 72.0%; Score 36; DB 1; Length 111;  
 Best Local Similarity 77.8%; Pred. No. 4.4;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
 : I I I I I I  
 Db 93 QHSRELPLT 101

RESULT 12  
 KVMS85  
 Ig kappa chain V regions (PC2485, PC4039) - mouse (tentative sequence)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 31-Mar-2000  
 C:Accession: A01939  
 R:Weigert, M.; Gatmaitan, L.; Loh, E.; Schilling, J.; Hood, L.  
 Nature 276, 785-790, 1978  
 A:Title: Rearrangement of genetic information may produce immunoglobulin diversity.  
 A:Reference number: A93204; MUID:79073152  
 A:Accession: A01939  
 A:Molecule type: protein  
 A:Residues: 1-111 <WEI>  
 A:Note: the PC4285 and PC4039 sequences are identical  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:16-94/Domain: immunoglobulin homology <IMM>  
 F:23-92/Disulfide bonds: #status predicted

Query Match 72.0%; Score 36; DB 1; Length 111;  
 Best Local Similarity 77.8%; Pred. No. 4.4;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
 : I I I I I I  
 Db 93 QHSRELPLT 101

RESULT 13  
 S06562

finger protein (clone XlCGF9-1) - African clawed frog (fragment)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 31-Dec-1993  
 C:Accession: S06562  
 R:Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeting, A.; Knoechel, J. Mol. Biol. 208, 639-659, 1989  
 A:Title: Second-order repeats in Xenopus laevis finger proteins.  
 A:Reference number: S05632; MUID:90040698  
 A:Accession: S06562  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-112 <NIE>  
 C:Keywords: DNA binding; zinc finger

Query Match 72.0%; Score 36; DB 2; Length 112;  
 Best Local Similarity 75.0%; Pred. No. 4.4;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
 I:| | | |  
 Db 56 HTGEKPFT 63

## RESULT 14

S06571  
 finger protein (clone XlCGF44-2) - African clawed frog (fragment)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 10-Oct-1997  
 C:Accession: S06571  
 R:Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeting, A.; Knoechel, J. Mol. Biol. 208, 639-659, 1989  
 A:Title: Second-order repeats in Xenopus laevis finger proteins.  
 A:Reference number: S05632; MUID:90040698  
 A:Accession: S06571  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-138 <NIE>  
 C:Superfamily: LIM metal-binding repeat homology  
 C:Keywords: DNA binding; zinc finger  
 F:34-93/Domain: LIM metal-binding repeat homology <LIM4>

Query Match 72.0%; Score 36; DB 2; Length 138;  
 Best Local Similarity 75.0%; Pred. No. 5.6;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
 I:| | | |  
 Db 54 HTGEKPFT 61

## RESULT 15

S06574  
 finger protein (clone XlCGF49-1) - African clawed frog (fragment)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 31-Dec-1993  
 C:Accession: S06574  
 R:Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeting, A.; Knoechel, J. Mol. Biol. 208, 639-659, 1989  
 A:Title: Second-order repeats in Xenopus laevis finger proteins.  
 A:Reference number: S05632; MUID:90040698  
 A:Accession: S06574  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-140 <NIE>  
 C:Keywords: DNA binding; zinc finger

Query Match 72.0%; Score 36; DB 2; Length 140;  
 Best Local Similarity 75.0%; Pred. No. 5.6;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
 I:| | | |  
 Db 28 HTGEKPFT 35

Search completed: June 28, 2001, 15:58:47  
 Job time: 375 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 15:54:39 ; Search time 105.36 Seconds  
(without alignments)  
2.926 Million cell updates/sec

Title: US-09-724-406-32  
Perfect score: 50  
Sequence: 1 QHSGELPFT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	39	78.0	223	1 ZG52_XENLA	P18727 xenopus lae
2	36	72.0	111	1 KV3S_MOUSE	P01671 mus musculus
3	36	72.0	111	1 KV3U_MOUSE	P01673 mus musculus
4	36	72.0	112	1 ZG9_XENLA	P18738 xenopus lae
5	36	72.0	138	1 ZG44_XENLA	P18721 xenopus lae
6	36	72.0	140	1 ZG49_XENLA	P18724 xenopus lae
7	36	72.0	157	1 Z02_XENLA	P18743 xenopus lae
8	36	72.0	169	1 ZG62_XENLA	P18731 xenopus lae
9	36	72.0	196	1 ZG3_XENLA	P18718 xenopus lae
10	36	72.0	196	1 ZG64_XENLA	P18732 xenopus lae
11	36	72.0	196	1 ZG8_XENLA	P18737 xenopus lae
12	36	72.0	196	1 Z036_XENLA	P18746 xenopus lae
13	36	72.0	214	1 Z010_XENLA	P18739 xenopus lae
14	36	72.0	247	1 Z020_XENLA	P18744 xenopus lae
15	36	72.0	252	1 ZG28_XENLA	P18716 xenopus lae
16	36	72.0	280	1 ZG46_XENLA	P18722 xenopus lae
17	36	72.0	336	1 ZG57_XENLA	P18729 xenopus lae
18	36	72.0	337	1 ZG26_XENLA	P18715 xenopus lae
19	36	72.0	349	1 NTRB_SCOLI	P08712 escherichia
20	36	72.0	349	1 NTRB_KLEPN	P06218 klebsiella
21	36	72.0	439	1 Z028_XENLA	P18747 xenopus lae
22	36	72.0	445	1 ZG5_XENLA	P18725 xenopus lae
23	36	72.0	453	1 Z06_XENLA	P18749 xenopus lae
24	36	72.0	474	1 Z141_HUMAN	O15928 homo sapien
25	36	72.0	647	1 ZG48_XENLA	P18723 xenopus lae
26	36	72.0	648	1 Z202_HUMAN	O95125 homo sapien
27	36	72.0	675	1 ZG20_XENLA	P18714 xenopus lae
28	36	72.0	697	1 Y441_HUMAN	O43167 homo sapien
29	35	70.0	321	1 Z177_HUMAN	O13360 homo sapien
30	35	70.0	406	1 ZF59_MOUSE	P16373 mus musculus
31	35	70.0	589	1 Z132_HUMAN	P52740 homo sapien
32	34	68.0	56	1 ZN26_HUMAN	P17031 homo sapien
33	34	68.0	108	1 KV3V_MOUSE	P01674 mus musculus

34	34	68.0	417	1 Z232_HUMAN	Q9uny5 homo sapien
35	34	68.0	471	1 VCP_AEDAE	P42560 aedes aegypt
36	34	68.0	488	1 ZF92_MOUSE	Q62396 mus musculus
37	34	68.0	512	1 Z169_HUMAN	Q14929 homo sapien
38	34	68.0	517	1 IKAR_MOUSE	Q03267 mus musculus
39	34	68.0	518	1 IKAR_CHICK	O42410 gallus gall
40	34	68.0	519	1 IKAR_HUMAN	Q13422 homo sapien
41	34	68.0	521	1 ZN35_HUMAN	P13682 homo sapien
42	34	68.0	522	1 IKAR_ONCMV	O13089 oncorhynchus
43	34	68.0	526	1 HELI_HUMAN	Q9uks7 homo sapien
44	34	68.0	526	1 HELI_MOUSE	P81183 mus musculus
45	34	68.0	537	1 Z029_XENLA	P18748 xenopus lae

## ALIGNMENTS

```

RESULT 1
ZG52_XENLA
ID ZG52_XENLA STANDARD; PRT; 223 AA.
AC P18727;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE GASTRULA ZINC FINGER PROTEIN XLCGF52.1 (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90040698; PubMed=2509712;
RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
RA Poetling A., Knoechel W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. Mol. Biol. 208:639-659(1989).
PIR: S06576; S06576.
DR HSSP; P25490; 12NM.
DR InterPro; IPR000822;
DR Pfam; PF000096; zf-C2H2; 8.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 8.
KW zinc_finger; Metal-binding; DNA-binding; Repeat.
FT NON_TER 1
FT ZN_FING 6 27 C2H2-TYPE.
FT ZN_FING 33 55 C2H2-TYPE.
FT ZN_FING 61 83 C2H2-TYPE.
FT ZN_FING 89 111 C2H2-TYPE.
FT ZN_FING 117 139 C2H2-TYPE.
FT ZN_FING 145 167 C2H2-TYPE.
FT ZN_FING 173 195 C2H2-TYPE.
FT ZN_FING 201 223 C2H2-TYPE.
FT NON_TER 223
SQ SEQUENCE 223 AA; 25549 MW; BFA977BB6701BA96 CRC64;

```

Query Match 78.0%; Score 39; DB 1; Length 223;  
Best Local Similarity 87.5%; Pred. No. 1.2;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HSGELPFT 9  
Db 195 HSGELPFT 202

```

RESULT 2
KV3S_MOUSE
ID KV3S_MOUSE STANDARD; PRT; 111 AA.
AC P01671;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

2

RL J. Mol. Biol. 208:639-659(1989).  
 DR PIR; S06571; S06571.  
 DR HSP; P08047; ISP2.  
 DR InterPro: IPR000822; -.  
 DR Pfam; PF00096; zf-C2H2; 5.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 5.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 5.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
 FT NON\_TER 1 1  
 FT ZN\_FING 5 27 C2H2-TYPE.  
 FT ZN\_FING 32 54 C2H2-TYPE.  
 FT ZN\_FING 60 82 C2H2-TYPE.  
 FT ZN\_FING 88 110 C2H2-TYPE.  
 FT ZN\_FING 116 138 C2H2-TYPE.  
 FT NON\_TER 138 138  
 SQ SEQUENCE 138 AA; 16316 MW; 3FC025908EEF43AB CRC64;

Query Match 72.0%; Score 36; DB 1; Length 138;  
 Best Local Similarity 75.0%; Pred. No. 2.8;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPPT 9  
 DB 54 HTGKPT 61

RESULT 6  
 ZG49\_XENLA STANDARD; PRT; 140 AA.  
 ID ZG49\_XENLA  
 AC P18724;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE GASTRULA ZINC FINGER PROTEIN XLCGF49.1 (FRAGMENT).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90040698; PubMed=2509712;  
 RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,  
 RA Poeting A., Knoechel W.;  
 RT "Second-order repeats in Xenopus laevis finger proteins.";  
 RL J. Mol. Biol. 208:639-659(1989).  
 DR PIR; S06574; S06574.  
 DR HSP; P25490; 12NM.  
 DR InterPro: IPR000822; -.  
 DR Pfam; PF00096; zf-C2H2; 5.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 5.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 5.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
 FT NON\_TER 1 1  
 FT ZN\_FING 6 28 C2H2-TYPE.  
 FT ZN\_FING 34 56 C2H2-TYPE.  
 FT ZN\_FING 62 84 C2H2-TYPE.  
 FT ZN\_FING 90 112 C2H2-TYPE.  
 FT ZN\_FING 118 140 C2H2-TYPE.  
 FT NON\_TER 140 140  
 SQ SEQUENCE 140 AA; 16386 MW; EPA279A81583E277 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 140;  
 Best Local Similarity 75.0%; Pred. No. 2.9;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPPT 9  
 DB 28 HTGKPT 35

RESULT 7  
 ZG62\_XENLA STANDARD; PRT; 157 AA.  
 ID ZG62\_XENLA  
 AC P18743;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE OOCYTE ZINC FINGER PROTEIN XLCGF2 (FRAGMENT).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90040698; PubMed=2509712;  
 RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,  
 RA Poeting A., Knoechel W.;  
 RT "Second-order repeats in Xenopus laevis finger proteins.";  
 RL J. Mol. Biol. 208:639-659(1989).  
 DR PIR; S05632; S05632.  
 DR HSP; P08047; ISP2.  
 DR InterPro: IPR000822; -.  
 DR Pfam; PF00096; zf-C2H2; 5.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 5.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
 FT NON\_TER 1 1  
 FT ZN\_FING 6 28 C2H2-TYPE.  
 FT ZN\_FING 34 56 C2H2-TYPE.  
 FT ZN\_FING 79 101 C2H2-TYPE.  
 FT ZN\_FING 107 129 C2H2-TYPE.  
 FT ZN\_FING 135 157 C2H2-TYPE.  
 FT NON\_TER 157 157  
 SQ SEQUENCE 157 AA; 17953 MW; 5386C35255E9AA6 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 157;  
 Best Local Similarity 75.0%; Pred. No. 3.2;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPPT 9  
 DB 129 HTGKPT 135

RESULT 8  
 ZG62\_XENLA STANDARD; PRT; 169 AA.  
 ID ZG62\_XENLA  
 AC P18731;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE GASTRULA ZINC FINGER PROTEIN XLCGF62.1 (FRAGMENT).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90040698; PubMed=2509712;  
 RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,  
 RA Poeting A., Knoechel W.;  
 RT "Second-order repeats in Xenopus laevis finger proteins.";  
 RL J. Mol. Biol. 208:639-659(1989).  
 DR PIR; S06580; S06580.  
 DR HSP; P08047; ISP2.  
 DR InterPro: IPR000822; -.  
 DR Pfam; PF00096; zf-C2H2; 6.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 6.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 6.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.

```

FT NON_TER 1 1 C2H2-TYPE.
FT ZN_FING 6 28 C2H2-TYPE.
FT ZN_FING 34 56 C2H2-TYPE.
FT ZN_FING 62 84 C2H2-TYPE.
FT ZN_FING 90 113 C2H2-TYPE.
FT ZN_FING 119 141 C2H2-TYPE.
FT ZN_FING 147 169 C2H2-TYPE.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 19525 MW; 94EB7CBA8B177297 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 169;
Best Local Similarity 75.0%; Pred. No. 3.5;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9
DB 84 HTGEKPT 91

RESULT 9
ZG3_XENLA STANDARD; PRT; 196 AA.
AC P18718;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE GASTRULA ZINC FINGER PROTEIN XLCGF3.1 (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90040698; PubMed=2509712;
RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
RA Poetling A., Knoechel W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. Mol. Biol. 208:639-659(1989).
DR PIR; S06558; S06558.
DR HSP; P25490; 12NM.
DR InterPro; IPR000822;
DR Pfam; PF00096; zf-C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 7.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT NON_TER 1 1 C2H2-TYPE.
FT ZN_FING 6 28 C2H2-TYPE.
FT ZN_FING 34 56 C2H2-TYPE.
FT ZN_FING 62 84 C2H2-TYPE.
FT ZN_FING 90 112 C2H2-TYPE.
FT ZN_FING 118 140 C2H2-TYPE.
FT ZN_FING 146 168 C2H2-TYPE.
FT ZN_FING 174 196 C2H2-TYPE.
FT NON_TER 196 196
SQ SEQUENCE 196 AA; 22378 MW; B5735306E722FDC CRC64;

Query Match 72.0%; Score 36; DB 1; Length 196;
Best Local Similarity 75.0%; Pred. No. 4.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9
DB 28 HTGEKPT 35

RESULT 10
ZG64_XENLA STANDARD; PRT; 196 AA.
AC P18732;
DT 01-NOV-1990 (Rel. 16, Created)

```

```

DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE GASTRULA ZINC FINGER PROTEIN XLCGF4.1 (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90040698; PubMed=2509712;
RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
RA Poetling A., Knoechel W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. Mol. Biol. 208:639-659(1989).
DR PIR; S06581; S06581.
DR HSP; P08047; 1SP2.
DR InterPro; IPR000822;
DR Pfam; PF00096; zf-C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 7.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT NON_TER 1 1 C2H2-TYPE.
FT ZN_FING 6 28 C2H2-TYPE.
FT ZN_FING 34 56 C2H2-TYPE.
FT ZN_FING 62 84 C2H2-TYPE.
FT ZN_FING 90 112 C2H2-TYPE.
FT ZN_FING 118 140 C2H2-TYPE.
FT ZN_FING 146 168 C2H2-TYPE.
FT ZN_FING 174 196 C2H2-TYPE.
FT NON_TER 196 196
SQ SEQUENCE 196 AA; 23052 MW; 15B0E600C8FEF9A8 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 196;
Best Local Similarity 75.0%; Pred. No. 4.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9
DB 84 HTGEKPT 91

RESULT 11
ZG8_XENLA STANDARD; PRT; 196 AA.
AC P18737;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE GASTRULA ZINC FINGER PROTEIN XLCGF8.2DB (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90040698; PubMed=2509712;
RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
RA Poetling A., Knoechel W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. Mol. Biol. 208:639-659(1989).
DR PIR; S06561; S06561.
DR HSP; P25490; 12NM.
DR InterPro; IPR000822;
DR Pfam; PF00096; zf-C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 7.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT NON_TER 1 1 C2H2-TYPE.
FT ZN_FING 6 28 C2H2-TYPE.
FT ZN_FING 34 56 C2H2-TYPE.

```

FT ZN\_FING 62 84 C2H2-TYPE.  
 FT ZN\_FING 90 112 C2H2-TYPE.  
 FT ZN\_FING 118 140 C2H2-TYPE.  
 FT ZN\_FING 146 168 C2H2-TYPE.  
 FT ZN\_FING 174 196 C2H2-TYPE.  
 FT NON\_TER 196  
 SQ SEQUENCE 196 AA; C65153F709EDE461 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 196;  
 Best Local Similarity 75.0%; Pred. No. 4.1;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
 Db 84 HTGKPT 91

RESULT 12  
 ZO26\_XENLA STANDARD; PRT; 196 AA.  
 AC P18746;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE OOCYTE ZINC FINGER PROTEIN XLCOP26 (FRAGMENT).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90040698; PubMed=2509712;  
 RA Niefeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,  
 RA Poeting A., Knoechel W.;  
 RT "Second-order repeats in Xenopus laevis finger proteins.";  
 RL J. Mol. Biol. 208:639-659(1989).  
 DR PIR; S06555; S06553.  
 DR HSP; P25490; IZNM.  
 DR InterPro; IPR000822;  
 DR Pfam; PF00096; zf-C2H2; 7.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 7.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 7.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
 FT NON\_TER 1  
 FT ZN\_FING 6 28 C2H2-TYPE.  
 FT ZN\_FING 34 56 C2H2-TYPE.  
 FT ZN\_FING 62 84 C2H2-TYPE.  
 FT ZN\_FING 90 112 C2H2-TYPE.  
 FT ZN\_FING 118 140 C2H2-TYPE.  
 FT ZN\_FING 146 168 C2H2-TYPE.  
 FT ZN\_FING 174 196 C2H2-TYPE.  
 FT NON\_TER 196  
 SQ SEQUENCE 196 AA; 22659 MW; A4417A032257D595 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 196;  
 Best Local Similarity 75.0%; Pred. No. 4.1;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
 Db 84 HTGKPT 91

RESULT 13  
 ZO10\_XENLA STANDARD; PRT; 214 AA.  
 AC P18739;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)

DE OOCYTE ZINC FINGER PROTEIN XLCOP10 (FRAGMENT).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89005062; PubMed=3139407;  
 RA Koester M., Pieler T., Poeting A., Knoechel W.;  
 RT "The finger motif defines a multigene family represented in the  
 RT maternal mRNA of Xenopus laevis oocytes.";  
 RL EMBO J. 7:1735-1741(1988).  
 DR PIR; S00832; S00832.  
 DR HSP; P08047; ISP2.  
 DR InterPro; IPR000822;  
 DR Pfam; PF00096; zf-C2H2; 7.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 6.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 7.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
 FT NON\_TER 1  
 FT ZN\_FING 1 23 C2H2-TYPE.  
 FT ZN\_FING 29 51 C2H2-TYPE.  
 FT ZN\_FING 57 79 C2H2-TYPE.  
 FT ZN\_FING 85 107 C2H2-TYPE.  
 FT ZN\_FING 113 135 C2H2-TYPE.  
 FT ZN\_FING 141 163 C2H2-TYPE.  
 FT ZN\_FING 169 191 C2H2-TYPE.  
 SQ SEQUENCE 214 AA; 24830 MW; DCEEE00DEDED1A5F8 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 214;  
 Best Local Similarity 75.0%; Pred. No. 4.4;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
 Db 107 HTGKPT 114

RESULT 14  
 ZO20\_XENLA STANDARD; PRT; 247 AA.  
 AC P18744;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE OOCYTE ZINC FINGER PROTEIN XLCOP20 (FRAGMENT).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90040698; PubMed=2509712;  
 RA Niefeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,  
 RA Poeting A., Knoechel W.;  
 RT "Second-order repeats in Xenopus laevis finger proteins.";  
 RL J. Mol. Biol. 208:639-659(1989).  
 DR PIR; S06553; S06553.  
 DR HSP; P08047; ISP2.  
 DR InterPro; IPR000822;  
 DR Pfam; PF00096; zf-C2H2; 8.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 7.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 8.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
 FT NON\_TER 1  
 FT ZN\_FING 6 28 C2H2-TYPE.  
 FT ZN\_FING 34 56 C2H2-TYPE.  
 FT ZN\_FING 62 84 C2H2-TYPE.  
 FT ZN\_FING 90 112 C2H2-TYPE.  
 FT ZN\_FING 118 140 C2H2-TYPE.

FT ZN\_FING 146 168 C2H2-TYPE.  
FT ZN\_FING 174 196 C2H2-TYPE.  
FT ZN\_FING 225 247 C2H2-TYPE.  
FT NON\_TER 247 247  
SQ SEQUENCE 247 AA; 28276 MW; C76B7808F6BE09AD CRC64;

Query Match 72.0%; Score 36; DB 1; Length 247;  
Best Local Similarity 75.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
|:|:|  
Db 112 HTGEKPFT 119

RESULT 15  
ZG28\_XENLA STANDARD; PRT; 252 AA.  
AC PI8716;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE GASTRULA ZINC FINGER PROTEIN XLCGF28.1 (FRAGMENT).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-90040698; PubMed-2509712;  
RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,  
RA Poetig A., Knoechel W.;  
RT "Second-order repeats in Xenopus laevis finger proteins.";  
RL J. Mol. Biol. 208:639-659(1989).  
DR PIR; S06567; S06567.  
DR HSP; P08047; LSP2.  
DR InterPro; IPR000822;  
DR Pfam; PF00096; zf-C2H2; 9.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 9.  
DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 9.  
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
FT NON\_TER 1  
FT ZN\_FING 6 28 C2H2-TYPE.  
FT ZN\_FING 34 56 C2H2-TYPE.  
FT ZN\_FING 62 84 C2H2-TYPE.  
FT ZN\_FING 90 112 C2H2-TYPE.  
FT ZN\_FING 118 140 C2H2-TYPE.  
FT ZN\_FING 146 168 C2H2-TYPE.  
FT ZN\_FING 174 196 C2H2-TYPE.  
FT ZN\_FING 202 224 C2H2-TYPE.  
FT ZN\_FING 230 252 C2H2-TYPE.  
FT NON\_TER 252 252  
SQ SEQUENCE 252 AA; 28872 MW; 13A72370B10F0BD7 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 252;  
Best Local Similarity 75.0%; Pred. No. 5.2;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
|:|:|  
Db 28 HTGEKPFT 35

Search completed: June 28, 2001, 15:54:40  
Job time: 128 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2001, 16:08:27 ; Search time 411.58 Seconds  
(without alignments)  
2.893 Million cell updates/sec

Title: US-09-724-406-32  
Perfect score: 50  
Sequence: 1 QHSGELPFT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues  
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL16: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_unclassified: \*  
13: sp\_vertebrate: \*  
14: sp\_virus: \*

\*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	82.0	556	4 Q9UMP5	Q9ump5 homo sapien
2	40	80.0	488	10 Q9LGM4	Q9lqm4 oryza sativ
3	39	78.0	278	5 Q9V718	Q9v718 drosophila
4	39	78.0	405	4 Q9UFB7	Q9ufb7 homo sapien
5	39	78.0	571	4 Q9GZV8	Q9gzv8 homo sapien
6	39	78.0	594	5 Q20537	Q20537 caenorhabdi
7	39	78.0	606	4 Q9V2D9	Q9v2d9 homo sapien
8	37	74.0	181	5 Q9W3J3	Q9w3j3 drosophila
9	37	74.0	384	2 Q46239	Q46239 clostridium
10	37	74.0	498	4 Q9H8L4	Q9h8l4 homo sapien
11	37	74.0	501	5 Q9VG72	Q9vg72 drosophila
12	37	74.0	525	5 Q9WN56	Q9wn56 drosophila
13	37	74.0	532	5 Q01960	Q01960 caenorhabdi
14	37	74.0	720	6 Q9N003	Q9n003 macaca fasc
15	37	74.0	756	5 Q61362	Q61362 drosophila
16	37	74.0	891	5 Q61361	Q61361 drosophila
17	37	74.0	934	5 Q9WKF1	Q9wkf1 drosophila
18	37	74.0	936	5 Q9VGK1	Q9vgk1 drosophila
19	37	74.0	962	5 Q61360	Q61360 drosophila

20	37	74.0	1143	10 Q92PS9	Q9zps9 arabidopsis
21	36	72.0	52	4 Q15923	Q15923 homo sapien
22	36	72.0	90	4 Q92669	Q92669 homo sapien
23	36	72.0	306	4 Q9NQJ3	Q9nqj3 homo sapien
24	36	72.0	349	2 Q9ZH35	Q9zh35 enterobacte
25	36	72.0	354	13 Q09047	Q09047 xenopus lae
26	36	72.0	382	5 Q9VP31	Q9vp31 drosophila
27	36	72.0	420	13 Q9PSM1	Q9psm1 xenopus lae
28	36	72.0	462	13 Q91906	Q91906 xenopus lae
29	36	72.0	478	4 Q9H5P0	Q9h5p0 homo sapien
30	36	72.0	484	4 Q13397	Q13397 homo sapien
31	36	72.0	493	13 Q08701	Q08701 xenopus lae
32	36	72.0	522	4 Q99592	Q99592 homo sapien
33	36	72.0	522	11 Q9WUK6	Q9wuk6 mus musculu
34	36	72.0	522	11 Q9JKY3	Q9jky3 rattus norv
35	36	72.0	547	4 Q9V5A5	Q9v5a5 homo sapien
36	36	72.0	549	5 Q9VE54	Q9ve54 drosophila
37	36	72.0	556	5 Q9VDN4	Q9vnd4 drosophila
38	36	72.0	556	5 Q9GRA5	Q9gra5 halocynthia
39	36	72.0	591	13 Q09046	Q09046 xenopus lae
40	36	72.0	634	4 Q9NSM4	Q9nsm4 homo sapien
41	36	72.0	648	4 Q9H1B9	Q9h1b9 homo sapien
42	36	72.0	692	4 Q9UJW8	Q9ujw8 homo sapien
43	36	72.0	692	4 Q9PIU2	Q9piu2 homo sapien
44	36	72.0	1001	4 Q9NND7	Q9nnd7 homo sapien
45	36	72.0	1239	5 Q9VT48	Q9vt48 drosophila

## ALIGNMENTS

RESULT 1  
Q9UMP5  
ID AC Q9UMP5; PRELIMINARY; PRT; 556 AA.  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-MAR-2001 (TREMREL. 16, Last annotation update)  
DE AP4 PROTEIN.  
GN AP4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dobner T.G., Fischer M., Groitl P.;  
RT "Cloning of a novel zinc finger protein."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ245553; CAB53100.1; -  
DR HSSP; P08047; ISP2.  
DR InterPro; IPR000822; -  
DR Pfam; PF00096; zf-C2H2; 15.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 14.  
DR SMART; SM00355; ZnF\_C2H2; 1.  
KW DNA-binding; Metal-binding; Zinc-finger.  
SQ SEQUENCE 556 AA; 62286 MW; E9CFFB4AF/CL2E90 CRC64;

Query Match 82.0%; Score 41; DB 4; Length 556;  
Best Local Similarity 77.8%; Pred No. 4;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
Db 385 QHTGERPFT 393  
||:|:| |

RESULT 2  
Q9LGM4  
ID Q9LGM4 PRELIMINARY; PRT; 488 AA.  
AC Q9LGM4;  
DT 01-OCT-2000 (TREMREL. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE ESTS AU075348(C11252).  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;  
 CC Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0041E11".  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0433F09".  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP002521; BAA96760.1; -.  
 DR EMBL; AP002539; BAB08199.1; -.  
 DR InterPro; IPR001841; -.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 SQ SEQUENCE 488 AA; 53169 MW; 1FBAD40A7463C323 CRC64;

Query Match 80.0%; Score 40; DB 10; Length 488;  
 Best Local Similarity 75.0%; Pred. No. 5.6;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HSGELPPT 9  
 ||||:|:|  
 Db 235 HSGEMPPT 242

RESULT 3  
 Q9V718 ID Q9V718 PRELIMINARY; PRT; 278 AA.  
 AC Q9V718;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE CG17385 PROTEIN.  
 GN CG17385.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aubayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadenot E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AF003814; AAF58248.1; -.  
 DR HSP; P08047; 1SP2  
 DR Flybase; FBgn0033934; CG17385.  
 DR InterPro; IPR000822; -.  
 DR Pfam; PF00096; zf-C2H2; 7.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 7.  
 DR SMART; SM00355; ZnF\_C2H2; 1.  
 KW DNA-binding; Metal-binding; Zinc-finger.  
 SQ SEQUENCE 278 AA; 32463 MW; 3D6DEIAE5D77C223 CRC64;

Query Match 78.0%; Score 39; DB 5; Length 278;  
 Best Local Similarity 87.5%; Pred. No. 4.9;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HSGELPPT 9  
 |||||  
 Db 96 HSGERPPT 103

RESULT 4  
 Q9UFB7 ID Q9UFB7 PRELIMINARY; PRT; 405 AA.  
 AC Q9UFB7;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE HYPOPHETICAL 46.0 KDA PROTEIN (FRAGMENT).  
 GN DKFZP434N0615.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RA Bloeker H., Boecker M., Brandt P., Mewes H.W., Gassenhuber J.,  
 RA Wiemann S.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -|- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL; AL133062; CAB61386.1; -.  
 DR HSP; P08047; 1SP2.  
 DR InterPro; IPR000822; -.  
 DR Pfam; PF00096; zf-C2H2; 9.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 7.  
 DR SMART; SM00355; ZnF\_C2H2; 1.  
 KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;  
 ZINC-finger.  
 FT NON-TER  
 SQ SEQUENCE 405 AA; 46007 MW; 72116A2A7ACDCB33 CRC64;

Query Match 78.0% Score 39; DB 4; Length 405;  
 Best Local Similarity 87.5%; Pred. No. 7.3;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPF 8  
 Db 227 QHSGEKPF 234  
 RESULT 5  
 Q9GV8 PRELIMINARY; PRT; 571 AA.  
 AC Q9GV8;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE CONA FLJ12533 FIS, CLONE NT2RM4000202, WEAKLY SIMILAR TO ZINC FINGER  
 DE PROTEIN MOK-2 (PR-DOMAIN CONTAINING PROTEIN 14).  
 GN PRD14.  
 OS Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Wagatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takanashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Du Y., Yang X., Huang S.;  
 RT "A family of novel PR-domain (PRDM) genes as candidate tumor  
 RT suppressors."  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK022595; BAB14120.1; -;  
 DR EMBL; AF319458; AAG39635.1; -;  
 SQ SEQUENCE 571 AA; 64062 MW; 60D5ACC96DADE500 CRC64;

Query Match 78.0% Score 39; DB 4; Length 571;  
 Best Local Similarity 87.5%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPF 8  
 Db 510 QHSGEKPF 517  
 RESULT 6  
 Q20537 PRELIMINARY; PRT; 594 AA.  
 AC Q20537;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE SIMILAR TO C2H2-TYPE ZINC FINGER.  
 GN F47E1.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Feloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lighning J., L.oyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans."  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Favellio T.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL; U39993; AAA81086.1; -;  
 DR HSP; P08046; 1A1H.  
 DR InterPro: IPR000822; -;  
 DR Pfam: PF00096; zf-C2H2; 4.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR PROSITE; PS00026; ZINC\_FINGER\_C2H2; 3.  
 DR SMART; SM00355; Znf\_C2H2; 1.  
 KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.  
 SQ SEQUENCE 594 AA; 67531 MW; 81957E08C0CB5A7 CRC64;

Query Match 78.0% Score 39; DB 5; Length 594;  
 Best Local Similarity 87.5%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPF 8  
 Db 573 QHSGEKPF 580  
 RESULT 7  
 Q9Y2D9 PRELIMINARY; PRT; 606 AA.  
 AC Q9Y2D9;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE KIAA0924 PROTEIN.  
 GN KIAA0924.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRIN;  
 RX MEDLINE=99246063; PubMed=10231032;  
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirokawa M.,  
 RA Miyajima N., Tanaka A., Kofani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 6:63-70(1999).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL; AB023141; BAA76768.1; -;  
 DR HSP; P08047; 1SP2.  
 DR InterPro: IPR000522; -;  
 DR Pfam; PF00096; zf-C2H2; 9.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 7.

DR SMART: SM00355; ZnF\_C2H2; 1.  
 KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.  
 SQ SEQUENCE 606 AA; 69744 MW; C066D10E179985CC CRC64;

Query Match 78.0%; Score 39; DB 4; Length 606;  
 Best Local Similarity 87.5%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPFP 8  
 |||||  
 Db 378 QHSGEKPF 385

RESULT 8  
 Q9W3J3 PRELIMINARY; PRT; 181 AA.  
 AC Q9W3J3;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE CGL15336 PROTEIN.  
 GN CGL15336.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Cocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Heit J.G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RL EMBL; AE003443; AAF4634.1;  
 DR FlyBase; FBgn0030009; CGL15336.  
 DR InterPro; IPR000822; -.

DR Pfam: PF00096; zf-C2H2; 4.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 3.  
 DR SMART: SM00355; ZnF\_C2H2; 1.  
 KW DNA-binding; Metal-binding; Zinc-finger.  
 SQ SEQUENCE 181 AA; 21222 MW; E6169EDF0025D605 CRC64;

Query Match 74.0%; Score 37; DB 5; Length 181;  
 Best Local Similarity 55.6%; Pred. No. 7.9;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QHSGELPFP 9  
 :||:|:  
 Db 58 RUTGEMPF 66

RESULT 9  
 Q46239 PRELIMINARY; PRT; 384 AA.  
 AC Q46239;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE NANH GENE & ORF1.2,3 & 4 PRECURSOR.  
 OS Clostridium perfringens.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A99;  
 RX MEDLINE=95102306; PubMed=7804004;  
 RA Traving C., Schauer R., Roggentin P.;  
 RT "Gene structure of the 'large' sialidase isoenzyme from Clostridium  
 perfringens A99 and its relationship with other clostridial nanH  
 proteins.";  
 RL Glycoconj. J. 11:141-151(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A99;  
 RA Schauer R.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X87369; CAA60797.1; -.  
 DR InterPro; IPR001798; -.  
 DR Pfam; PF01344; Kelch; 2.  
 KW Signal.  
 FT SIGNAL. 1 19 POTENTIAL.  
 FT CHAIN 20 384  
 SQ SEQUENCE 384 AA; 42083 MW; E1E2589AFB1C6A9 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 384;  
 Best Local Similarity 66.7%; Pred. No. 18;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QHSGELPFP 9  
 :|:|:  
 Db 157 KHIGDLPT 165

RESULT 10  
 Q9H8L4 PRELIMINARY; PRT; 498 AA.  
 AC Q9H8L4;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE CDNA FLJ13479 FIS, CLONE PLACE1003738, WEAKLY SIMILAR TO ZINC FINGER  
 DE PROTEIN 135.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

```

RN RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RA Isogal T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masubo Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK023541; BABI4602.1; -.
SQ SEQUENCE 498 AA; 54872 MW; C704FCF55CCDAE1A CRC64;

Query Match 74.0%; Score 37; DB 4; Length 498;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HSGELPF 8
Db 13 HAGELPF 19

RESULT 11
Q9VG72
ID Q9VG72 PRELIMINARY; PRT; 501 AA.
AC Q9VG72;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CG5245 PROTEIN.
GN CG5245.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Boldwin D.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reineb K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

```

```

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003696; AAF54813.1; -.
DR HSSP; P08047; 1SP2.
DR FlyBase; FBgn0038047; CG5245.
DR InterPro; IPR000822; -.
DR Pfam; PF00096; zfc2H2; 15.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 14.
DR SMART; SM00355; Znf_C2H2; 1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 501 AA; 58790 MW; BBAE04740D41C43F CRC64;

Query Match 74.0%; Score 37; DB 5; Length 501;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HSGELPF 8
Db 438 HTGELPF 444

RESULT 12
Q9VN56
ID Q9VN56 PRELIMINARY; PRT; 525 AA.
AC Q9VN56;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CG14655 PROTEIN.
GN CG14655.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Boldwin D.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reineb K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

```

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasaman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003605; AAF52094.1; -;  
DR HSSP; P07248; IPAA.  
DR FlyBase; FBgn0037275; CG14655.  
DR InterPro; IPR000822; -;  
DR Pfam; PF00096; zf-C2H2; 7.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 7.  
DR SMART; SM00355; ZNF\_C2H2; 1.  
DR DNA-binding; Metal-binding; Zinc-finger.  
KW SEQUENCE 525 AA; 58076 MW; 69DEIAE3CA739727 CRC64;  
SQ

Query Match 74.0%; Score 37; DB 5; Length 525;  
Best Local Similarity 62.5%; Pred. No. 25;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 HSGELPFT 9  
|:|:|:|:  
Db 288 HTGEMPT 295

RESULT 13  
O01960 PRELIMINARY; PRT; 532 AA.  
AC O01960;  
DT 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE SIMILARITY TO MULTIPLE C2H2-TYPE ZINC FINGERS.  
GN C27A12.3.  
OS *Caenorhabditis elegans*.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Smaildon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,  
RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,  
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*.";  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN-BRISTOL N2;  
RA Murray J., Wohlmann P., Biewald T.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN-BRISTOL N2;  
RA Waterston R.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX STRAIN-BRISTOL N2;  
RA Waterston R.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
DR EMBL; AF003137; AAB93641.1; -;  
DR HSSP; P07248; 2ADR.  
DR InterPro; IPR000822; -;  
DR Pfam; PF00096; zf-C2H2; 6.  
DR PRINTS; PR00048; ZINC\_FINGER.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; UNKNOWN\_6.  
DR SMART; SM00355; ZNF\_C2H2; 1.  
DR DNA-binding; Nuclear protein; Zinc-finger.  
KW SEQUENCE 532 AA; 61223 MW; 154682344C6D2E4E CRC64;  
SQ

Query Match 74.0%; Score 37; DB 5; Length 532;  
Best Local Similarity 75.0%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 8  
|:|:|:|:  
Db 113 QHMGDLPT 120

RESULT 14  
Q9N003 PRELIMINARY; PRT; 720 AA.  
AC Q9N003;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE HYPOTHETICAL 83.9 KDA PROTEIN.  
OS *Macaca fascicularis* (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=CEREBELLUM CORTEX;  
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
RA Suzuki Y., Sugano S., Hashimoto K.;  
RT "Isolation of full-length cDNA clones from macaque brain cDNA  
RT libraries.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB046644; BAB03562.1; -;  
DR InterPro; IPR000822; -;  
DR InterPro; IPR001909; -;  
DR Pfam; PF00096; zf-C2H2; 19.  
DR PROSITE; PS50805; KRAB; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 19.  
DR SMART; SM00349; KRAB; 1.  
KW DNA-binding; Hypothetical protein; Metal-binding; Zinc-finger.  
SQ SEQUENCE 720 AA; 83922 MW; 074A7736A5903D90 CRC64;  
SQ

Query Match 74.0%; Score 37; DB 6; Length 720;  
Best Local Similarity 66.7%; Pred. No. 35;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QHSGELPFT 9  
|:|:|:|:  
Db 599 QHSGCKPFTS 607

RESULT 15  
O61362 PRELIMINARY; PRT; 756 AA.  
AC O61362;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE CROL GAMMA.  
GN CROL OR CROL GAMMA OR CG14938.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_taxid=7227;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON S;  
 RA D'Avino P.P., Thummel C.S.;  
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL; AF020349; AAC15518.1; -.  
 DR HSP; P08047; ISP2.  
 DR FlyBase; FBgn020309; crol.  
 DR InterPro; IPR000822; -.  
 DR Pfam; PF00096; zf-C2H2; 12.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 12.  
 DR SMART; SM00355; ZnF\_C2H2; 1.  
 KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.  
 SQ SEQUENCE 756 AA; 82883 MW; FA5DF25B09B74642 CRC64;

Query Match 74.0%; Score 37; DB 5; Length 756;  
 Best Local Similarity 75.0%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 QHSGELPF 8  
 Db 288 QHTGETPF 295  
 ||:|||||

Search completed: June 28, 2001, 16:08:28  
 Job time: 956 sec

